

Training procedural tasks through the manipulation of multimedia in dynamic visual displays with computer-based technologies.

James Kelly

School of Computing and Creative Technologies

University of Abertay Dundee

Training procedural tasks through the manipulation of multimedia in dynamic visual displays with computer-based technologies.

A thesis submitted in partial fulfilment of the requirements of the University of Abertay Dundee for the degree of Doctor of Philosophy

October 2004

I certify that this thesis is the true and accurate version of the thesis approved by the examiners.

Signed.....
(Director of studies.)

Date: 30th November 2004

Table of Contents

1	Introduction and Research Background.	10
1.1	Thesis Overview	10
1.2	Rationale behind research.....	11
1.3	Research Programme.....	12
1.4	Structure of the thesis.	14
1.5	Socio-economic background to study.	15
1.5.1	Shift from Fordist production system.	16
1.5.2	Agile manufacturing and manual assembly	18
1.5.3	Worker flexibility and training.	19
1.5.4	Just-in-time training and electronic support systems.....	21
1.6	Locus of Control.....	24
1.7	Chapter Summary	29
2	Head-mounted display technology.	31
2.1	Emerging technologies for industrial training.	31
2.2	Head-mounted display technology.	33
2.2.1	Cognitive Issues with head-mounted displays and head-up displays.	38
2.2.2	Previous Research with head-mounted displays.....	40
2.2.3	Previous research with head-up displays.....	43
2.3	Head-mounted display technology used in the first phase.	45
2.4	Chapter Summary	51
3	An Investigation into using head-mounted displays for following multimedia procedural instructions for an assembly task.	52
3.1	Phase 1 of experiments.....	52
3.1.1	Programme of experiments in phase 1.....	53
3.2	Experiment 1.1: Efficacy of HMD Compared to Desktop and Paper for Presenting Instructions for Assembling an Abstract Model.....	55
3.2.1	Method.....	58
3.2.2	Quantitative Results.....	65
3.2.3	Qualitative Results.....	66
3.2.4	Discussion.	70
3.3	Experiment 1.2: Efficiency Between two types of HMD and Desktop Instructions for Assembling an Abstract Model.....	73
3.3.1	Method.....	74
3.3.2	Quantitative Results.....	79
3.3.3	Qualitative Results.....	81
3.3.4	Discussion	84
3.4	Experiment 1.3: Comparison between two Modes of using a Head-mounted Display and Desktop Instructions for Assembling an Abstract Model.	86
3.4.1	Method.....	87
3.4.2	Quantitative Results.....	90
3.4.3	Qualitative Results.....	91
3.4.4	Discussion	95
3.5	General Conclusions for first phase of experiments.....	96

4	Focus of the second phase of research programme.	99
4.1	Factors that effect comparisons between head-mounted displays.	99
4.1.1	The existence of individual differences in Interpupillary Distances.	100
4.1.2	Problems associated with field of view and screen image.	105
4.2	Evolving design of head-mounted displays and wearable computers.	110
4.3	Chapter summary.	113
5	Introduction to second phase of experiments.	115
5.1	Rationale behind second phase of experiments.	115
5.2	The architecture of human memory.	116
5.2.1	Long-term memory store.	119
5.3	Two subtypes of long term memory.	120
5.3.1	Evidence for the locus of declarative and the locus of procedural memory.	122
5.3.2	Implicit memory and multiple memory systems.	126
5.4	Guidelines for multimedia learning.	129
5.5	Dual coding theory.	131
5.5.1	Research involving pictures and text.	134
5.6	Modifications of dual coding theory.	135
5.6.1	Cognitive load theory.	138
5.6.2	Theoretical framework of cognitive load theory.	140
5.6.3	Constructivism and computer based learning.	143
5.7	Principles for multimedia learning.	146
5.8	The multimedia paradigm.	149
5.8.1	Video technology in multimedia.	151
5.8.2	Sound in Multimedia.	152
5.8.3	Text in Multimedia.	155
5.9	Chapter summary.	158
6	Design and methodology of second phase of experiments.	161
6.1	Focus of second phase of experiments in research programme.	161
6.2	New experimental paradigm.	163
6.3	Previous Research on Assembly tasks using Model Kits.	167
6.3.1	Rationale behind research into video demonstrations.	169
6.3.2	Previous Research with Video Demonstrations.	174
6.3.3	Methodological differences between studies.	181
6.4	Taxonomy of Procedural tasks.	182
6.4.1	Declarative elements in procedural learning.	184
6.5	Hardware used in second phase of experiments.	187
6.5.1	Accommodation problems using Head-mounted displays.	187
6.6	Chapter Summary.	189
7	The Effectiveness of Different Multimedia Video Presentations for the Training and Retention of Procedural task.	192
7.1	Chapter overview.	192
7.2	Experiment 2.1: The Effectiveness of a Demonstration Video with Spoken Instructions for the Learning and Retention of an Assembly Task.	196
7.2.1	Method.	198
7.2.2	Quantitative Results.	207

7.2.3	Qualitative Results.....	210
7.2.4	Discussion	212
7.3	Experiment 2.2: A comparison between spoken instructions, text instructions, text plus spoken instructions and no text or spoken instructions on a demonstration video for the effectiveness of learning and retention of an assembly task.....	215
7.3.1	Method.....	216
7.3.2	Quantitative Results.....	221
7.3.3	Picture test.	227
7.3.4	Discussion	229
7.4	Experiment 2.3: The effect of a non verbal stimulus on a video demonstration for enhancing performance on an assembly task.....	231
7.4.1	Method.....	232
7.4.2	Quantitative Results.....	235
7.4.3	Discussion	238
7.5	Experiment 2.4: An investigation into the effectiveness of the contiguous presentation of voice instruction plus video demonstration compared to the separated presentation of voice instructions plus video demonstration for an assembly task.	239
7.5.1	Method.....	240
7.5.2	Results.	244
7.5.3	Discussion	246
7.6	General Discussion.....	248
8	Synthesis of findings, general conclusions and future research.	251
8.1	Chapter Overview.....	251
8.2	Aims of research.....	252
8.3	Conclusions drawn from the first phase of experiments.	254
8.4	Focus of second phase of experiments.	257
8.5	Conclusions drawn from the second phase of experiments.....	258
8.5.1	Results of the picture tests in experiments 2.1 and 2.2.....	262
8.6	Alternative research strategies for research.....	263
8.7	General conclusions and applications of research.....	265
8.8	Future research.	270
	References.....	274
	Appendices.....	301

List of Tables

Table 1.1: Comparison of traditional versus on-the-job training assumptions. (adapted from Najjar, Thompson & Ockerman(1999)).	23
Table 1.2: Traditional Training versus FAST (adapted from Ockerman, Najjar, Thompson & Atkinson (1996)).	25
Table 3.1: Programme of experiments for phase 1 of experiments.	53
Table 3. 2: Mean error rates and standard deviations (in brackets) for each instruction condition: Paper, Desktop and Head-Mounted Display.	66
Table 3.3: Qualitative feedback for Experiment 1.1	67
Table 3.4: Mean error rates and standard deviations (in brackets) for each instruction condition: Monitor, Glasstron and Desktop.	80
Table 3.5: Qualitative feedback for Experiment 1.2	82
Table 3. 6: Mean error rates and standard deviations (in brackets) for each instruction condition: Monitor, Glasstron and desktop.	91
Table 3. 7: Qualitative feedback for Experiment 1.3	92
Table 3.8: Summary of Quantitative results of experiments in Phase 1	96
Table 7.1: Programme of experiments for phase 2 of experiments.	193
Table 7.2: Maximum and Minimum Scores plus the Mean (and standard deviations in brackets) the Correctly named parts in the Picture Test.	212
Table 7. 3: Maximum and Minimum Scores plus the Mean (and standard deviations in brackets) the Correctly named parts in the Picture Test.	227
Table 7. 5: Summary of Quantitative results of experiments in Phase 2	248

List of Figures

Figure 2.1: Principle components of wearable computer developed by Georgia institute of technology. (Adapted from Thompson, Najjar & Ockermann (1997)).	32
Figure 2.2: Prototype wearable computer developed by Georgia institute of technology.(Adapted from Thopmson, Najjar & Ockermann (1997)).	34
Figure 2.3: Sony Glasstron Head-mounted display. (adapted from Sony instruction for the LDI-100BE 1999).	46
Figure 2.4: The Albatech Personal Monitor	48
Figure 3.1: Abstract models used in Experiment 1.1	60
Figure 3. 2:The interactive programme for Desktop and HMD conditions in Experiment 1.1	62
Figure 3.3:Testing bay with cameras used in experiment 1.1 Efficacy of HMD Compared to Desktop and Paper for Presenting Instructions for Assembling an Abstract Model.	63
Figure 3.4: Graph showing mean completion times and standard error in seconds for each instruction condition: Paper, Desktop and Head-Mounted Display.	65
Figure 3.5: Abstract models used in Experiment 2.1	75
Figure 3.6:Screenshot from Video clip used in Windows media player.	76
Figure 3. 7: Four way video analysis with CORIOscan	78
Figure 3.8: Graph Showing Mean Completion Times and Standard Error in seconds for each instruction condition, Albatech, Glasstron and Desktop.	79
Figure 3.9:Abstract Models used in Experiment 1.3	88
Figure 3.10: Graph showing mean completion times and standard error in seconds for each instruction condition: Glasstron Transparent, Glasstron Opaque and Desktop ..	90
Figure 4.1: Principal factors in Head-Mounted Display Design.	101
Figure 4. 2: Visual Distortion in Head-mounted Displays.	103
Figure 4.3: Liteye-D Amber Greyscale Microdisplay (Adapted from user manual 1999 and www.LYTEYE.com)	108
Figure 4.4: F.A.S.T. Wearable Computer System. Georgia Institute of Technology.	109
Figure 5. 1: Atkinson & Shiffrin (1968) multi-store model. Adapted from Baddeley (1987).	118
Figure 5.2: Brain regions associated with declarative memory. Adapted from	123
Figure 5.3: Brain areas associated with procedural memory. (Adapted from www.bvu.edu/faculty/ferguson/biotech/images/basalganglia)	125
Figure 5. 4: Schematic illustrating systems and brain structures involved in long-term memory. (Adapted from Squire & Zola-Morgan (1996)).	127
Figure 5.5:Schematic of Paivio's Dual Coding Theory (adapted from Paivio 1986).	132
Figure 5.6: The integrated dual-code hypothesis. (adapted from Mayer & Anderson 1991)	136
Figure 5.7: A dual coding model of multimedia learning. (adapted from Mayer & Sims 1994)	137

Figure 7.1: Model Car Used in Video for Pilot Study and Experiment 2.1	200
Figure 7.2: Script Used in Voice Narration for Pilot and Experiment 2.1	201
Figure 7.3: Instructions for Participants in the Experiments in the Second Phase. ...	203
Figure 7.4: Picture Test Given to Participants in Pilot and Experiment 2.1	204
Figure 7.5: Mean completion times and standard error in seconds in the initial training, criterion and assembly sessions for the video alone and the video with spoken instructions.	207
Figure 7.6: Mean number of times participants accessed the video clip in the initial and retention assembly sessions in the video only and the video plus narration conditions.....	209
Figure 7.7: Model of plane used in Experiment 2.2	219
Figure 7.8: script for Experiment 2.2.....	220
Figure 7. 9: Mean Difference in Completion Times for the four Conditions in the initial assembly.	221
Figure 7. 10: Mean completion times in the criterion assembly for the four conditions.	222
Figure 7.11: Mean completion times for the four conditions in the retention assembly.	223
Figure 7.12: mean video access frequency for the four conditions in the initial assembly.....	225
Figure 7.13: mean video access frequency for the four conditions in the retention assemblies.	226
Figure 7.14: Picture Test Given to Participants in Experiment 2.2	228
Figure 7.15: Model of helicopter used in video clip for experiment 2.3	234
Figure 7.16: Mean completion times and standard error in the initial, criterion and retention assemblies for the video only and the video plus white noise conditions. .	235
Figure 7.17: Mean number of times participants accessed the video clip in the Initial and Retention phases in the video only and the video plus white noise conditions. .	237
Figure 7. 18: Model truck used in experiment 2.4.....	242
Figure 7.19: Video script for experiment 2.4.....	243
Figure 7.20: Mean difference in completion times for the voice with sequential in the initial, criterion and retention assemblies.	244
Figure 7. 21: Mean difference in access frequency for the simultaneous and sequential voice conditions in the initial assembly and the retention assembly.	245

Acknowledgements.

Grateful thanks are due to the following:

Doctor Colin Cartwright, School of Computer and Advanced Technologies.
University of Abertay, Dundee.

Professor Michael Swanston, Vice principal of the University of Abertay Dundee.

Mr. Chris Milne, University of Abertay Dundee Library for his invaluable help with the Endnote referencing management software.

All the students and staff who participated in the experiments.

Mr. Ivicia Zriniski, research student at University of Abertay, Dundee, for his help with the hardware and software necessary for the production of the video clips used in the experiments.

Mrs. Eleanor Mollison, personal secretary to Professor Swanston.

Professor David Healy, School of Social Sciences. University of Abertay, Dundee.

Doctor Kenneth Scott-Brown, School of Social Sciences. University of Abertay, Dundee.

This work is dedicated to Hans Ebbinghaus and all the other pioneers of memory and learning research.

“The life so short, the craft so long to learn”.

Hippocrates (–357 BC)

Aphorisms, I. i (trans. Chaucer).

Abstract

This thesis set out to determine the optimum method of presenting multimedia instructional materials for learning a procedural task using a head-mounted display component of a wearable computer. There were four initial research questions; the first compared head-mounted display technology to traditional methods for following procedural task instructions. The second compared subtypes of head-mounted display for following task instructions and a third investigated possible differences between using an opaque and see-through head-mounted display for the task. A fourth research question was to involve exploratory experiments into the optimum configuration of text and vocal instructions on a video demonstration of the type used on a head-mounted display. The research programme comprised of two phases, the first phase investigated the first three questions. Results from the first phase indicated no significant difference in performing a procedural task on a head-mounted display in the conditions that investigated first and third research questions. Due to technological differences between the head-mounted displays the second research question could

not be addressed since comparisons between these displays were not feasible. This inability to make comparisons between subtypes resulted in second phase of experiments concentrating on the fourth research question; the investigation of the optimum presentation of multimedia learning materials on a video demonstration. The research in the second phase tested whether present guidelines for demonstrations for learning factual knowledge were applicable for creating multimedia demonstrations for learning procedural tasks. Results from the second phase of experiments challenge the use of the above guidelines for producing video demonstrations for learning a procedural task.

Chapter 1

1 Introduction and Research Background.

“While procedural tasks pervade the world of vocational education and training in the workplace, a serious problem is that students frequently do not retain much of what they have been taught about procedural tasks in school by the time they start working on the job. One way to counter practical skill loss is to develop instructional materials that are more resistant to forgetting and are more likely to promote transfer of skills to similar tasks and equipment.” (Ellis, Whitehill & Irick, 1996, p.130).

1.1 Thesis Overview

The above statement highlights the problem this study set out to investigate. Until comparatively recently the retention of procedural tasks in the workplace was largely unexplored by the developers of instructional material. The type of job or task, the amount and the quality of the training are factors that are postulated to contribute to poor retention (Farr, 1987). Another difficulty is that more often than not the equipment used to teach procedural tasks differs from the equipment used in the workplace. There are few studies in the literature that have sought to investigate the optimum way to develop instructional materials for procedural tasks. The majority of these studies investigated more traditional learning materials such as pictures and text rather than considering the possibilities of using multimedia instructions with new emerging computer-based technologies. Industrialists and researchers have become interested in harnessing the potential of the emerging technologies for training

procedural tasks in the workplace in an effort to make this training more effective.

Large, Behesti, Breuleaux, & Renaud, (1994) define a procedural task as a series of actions or steps that are executable by someone in order to achieve a goal. This study set out to address the problem of poor retention of workplace procedural training by attempting to ascertain the best use of an emerging technology and multimedia-training materials to learn a procedural task.

1.2 Rationale behind research.

The study was inspired by a growing interest in the mid to late 1990's for using emerging computer technology for training in the work environment. Shifts in industrial production systems in western societies in the last twenty to thirty years have made training an important issue in terms of cost and efficiency. Training workers quickly became a necessary practice since production lines and product types were constantly changing. The wearable computer was one emerging technology that was promoted as a potential aid for training in the workplace. A wearable computer is a mobile computing device that normally comprises of a headset containing a small screen or screens positioned near the eyes for viewing information originating from a battery powered mini microprocessor worn on the user's belt. The screens in the head-mounted display component may be see-through or opaque. The user controls the wearable computer by voice via a microphone or a handheld keyboard device. When coupled with the appropriate multimedia training materials the wearable computer was envisaged as a twenty first century solution to an ongoing skills training and retraining problem. The advantage of this technology is that it allows training on site, which is perceived as more efficient than using traditional off site training facilities.

This interest in using wearable computers coupled with multimedia training materials prompted several significant research questions that this study attempted to address. The nascent computer-based technology investigated in this study was the head-mounted display component or headset of a wearable computer. This thesis sought to investigate the best head-mounted display design coupled with the optimum configuration of multimedia training materials for learning a procedural task. The original intention of this research programme was to investigate four questions. The first was whether head-mounted displays were more efficient, less efficient or just the same as traditional methods for following instructions to perform a procedural task. A second question was whether there is a particular subtype of head-mounted display that is more efficient for following multimedia instructions. A third question investigated the possibility of a difference between a see-through head-mounted display and an opaque head-mounted display for following the instructions. A fourth and last question initially intended to conduct a preliminary investigation into the configuration of text, sound and moving image in the multimedia display used for presenting the procedural instructions in the head-mounted display. These questions were seen as important since research in the wearable computer literature has not fully considered human factor concerns with training using head-mounted displays and the importance of the configuration of the multimedia presentations for following and retaining procedural instructions.

1.3 Research Programme.

The research programme was conducted in two phases; the first investigated the first three of the above research questions. Due to technological differences between the head-mounted display subtypes being tested, it was not feasible to compare the

performance of head-mounted displays; this therefore limited this line of research. These difficulties in testing head-mounted displays led research in a second phase of experiments to focus on the fourth research question mainly using a desktop computer than a head-mounted display. This was due to the inability of the first phase of experiments to produce an optimum design and general usability problems with the head-mounted displays. The second phase developed into an investigation into the best configuration of vocal and text instructions on a video demonstration for learning and retaining a procedural task.

A body of research exists in the educational psychology literature that is concerned with the learning, retention and transfer of declarative knowledge, which entails facts, or cause and effect scientific explanations. This research dates back to the 1970s and has produced a number of theories or principles concerning the optimum presentation of multimedia or multi modal instructional materials for declarative knowledge. These principles have become over the years, guidelines or heuristics for courseware designers of computer-based learning materials that include both declarative (factual knowledge) and procedural knowledge. Some researchers however, for example (Michas & Berry 2000) have questioned the use of these guidelines for instructional materials for procedural tasks. This argument is based on the notion that declarative and procedural knowledge emanate from two distinct memory and learning systems.

The research undertaken in the second part of the thesis investigated the possibility that the above guidelines based on declarative learning for presenting multimedia instructions are not the same for procedural tasks. Another goal of this multimedia research was to extend previous research work carried out by researchers with more

traditional materials such as still pictures, illustrations and text captions. The starting point for establishing guidelines for multimedia research into procedural tasks in this thesis was to test whether the guidelines for declarative knowledge with animations were applicable for learning procedural tasks with video clips, since videos demonstrations are widely used in computer-based instruction.

1.4 Structure of the thesis.

The thesis comprises of two parts with each part covering one of the two phases of experiments. There are eight chapters; the first three are concerned with the research questions addressed in the first phase of experiments involving the head-mounted display component of a wearable computer. The first chapter gives an overview of the socio-economic background to this research and highlights the trend towards using emerging technology for training new tasks based on economic and efficiency reasons. This drive for efficient training may have human factors implications for the workers who must use these new training systems. This chapter highlights the importance of research into emerging technologies coupled with multimedia instructions for learning tasks. In chapter two there is an overview of the emerging technology that this thesis set out to investigate, the head-mounted display component of a wearable computer. The third chapter reports the results of the first phase of experiments. The first phase used the assembly of abstract Lego models to investigate the first three research questions. Chapter four describes the technological and usability problems encountered with comparing head-mounted displays with traditional methods and one another.

The fifth chapter introduces the second phase of experiments. This chapter explains the theory and current knowledge of memory function, since this is important for understanding the arguments for the existence of two memory systems and how this affects the optimum configuration of media for learning and retaining knowledge of tasks. Chapter six gives an overview of the development of multimedia learning materials in computer-based training in general. This sixth chapter outlines some of the limitations of earlier multimedia courseware that have recently been improved greatly by technological advances in computing. This has meant that there has been limited research in some areas of multimedia, especially sound and video. This chapter has a discussion about previous work in the area of the multimedia presentation of materials for learning the specific procedural task tested in the second phase, which was the assembly of models of transport vehicles from Lego. This chapter then describes the guidelines that will be tested in this thesis for learning procedural tasks. The seventh chapter reports the results of the second phase of experiments. The final chapter draws conclusions from the results of the experiments in the first phase. This chapter will then compare the results from the second phase to results from studies of learning declarative knowledge with multimedia. The findings from the second phase will be discussed in relation to methodological issues covered in chapters five and six of the thesis and give a general conclusion. At the end of this chapter there is a section that outlines possible further studies based on these results.

1.5 Socio-economic background to study.

This study set out to discover the most efficient head-mounted display design that when combined with a particular configuration of multimedia instructional materials

would best demonstrate a procedural task. This research theme became very relevant in the 1990s due to socio-economic changes in western industrial societies. In the 1990s researchers at the Georgia Institute of Technology tackled the problem of training and retraining of workers by investigating new technology. These researchers explored the possibility that emerging wearable computer technology coupled with the appropriate multimedia programmes could make training of the type of task encountered in modern industry more efficient. One particular paper written by the research team Ockerman, Najjar, Thompson, Treanor, & Atkinson (1996) at the Georgia Institute of Technology gives an insight into the prevailing zeitgeist. According to these authors the workplace of the nineties was in constant flux, largely due to massive upheavals that had occurred in the economies of advanced capitalist societies in the previous twenty years. There had been a shift away from a “Fordist” mass production practice and a gravitation towards more efficient production systems such as “Just-in time”. To understand why training became such an issue, it is important to chronicle these shifts in production methods.

1.5.1 Shift from Fordist production system.

Clarke (1990) describes “Fordism” as a term given to the mass production assembly line procedure and is named after the pioneer of this industrial production system, Henry Ford, the American car manufacturer. This production technique was pervasive for several decades in the twentieth century in advanced capitalist societies. The Fordist technique was based on the mass production of standardized products. This was achieved by employing the rigid technology of an assembly line with specialized machinery coupled with “Taylorist” work routines. The Taylorist working method as described by Ritzer (1992) is derived from the scientific management

technique created by F. W. Taylor at the beginning of the twentieth century. This constituted an important component of Fordism and entailed the precise detailing of each task on the assembly line down to the actions involved in each job and measuring the length of time these actions should take.

In the Fordist manufacturing system increased productivity was accomplished through economies of scale along with the deskilling, intensification and homogenisation of labour. These socio-economic conditions collectively created a cycle of rising living standards and increasing profits. After several decades of economic stability the Fordist society lapsed into crisis for a variety of reasons. Poire & Sable (1984) discuss the different socio-economic factors that led to the development of a crisis in the mass production economy. During the 1970s and 1980s there was social unrest, raw material shortages, rapid inflation combined with rising unemployment culminating in economic stagnation. Other historical events have been cited for the demise of the mass production society, these include; increases in oil prices, the abandonment by the United States government of its commitment to a fixed exchange rate between dollars and gold, and an economic downturn caused by an extended period of high interest rates in the United States.

In the view of Clarke (1990) this crisis of Fordist style mass production led to economic, social and political fragmentation out of which a new “Post-Fordist” era has appeared. As the Fordist production system reached its limits, new production techniques emerged, the saturation of mass markets led to a growing discrimination between products with the accent on style and quality. These differentiated products need briefer production runs so require shorter and more flexible production units

rather than a large assembly line. Industry looked to new emerging technologies to provide the method by which these flexible production systems could function profitably. However these new production systems have consequences for the workforce and industry. A more variable production system requires more highly skilled and adaptable workers.

1.5.2 Agile manufacturing and manual assembly

In a market where customisation mediates competitiveness the cost for redesigning automated processes for flexible production lines can be prohibitive, on the other hand workers are extremely adaptable. Manual assembly is commonplace in manufacturing processes where automation is not cost effective, products are customised or cannot be done by robots, for example the soldering of circuit boards. The early 1990's saw the dawn of a new manufacturing conceptual system that became known as "agile manufacturing". According to Newman et al.(2000) there are a number of definitions for agile manufacturing in the literature. It is often referred to as the ability to reconfigure a system quickly with a minimum of cost to produce different types of products.

The consequence of agile manufacturing is a mass customisation of small quantities of very specialised products depending to great extent on manual operations for flexibility. In a broader sense an agile manufacturing process can be defined as a manufacturing operation that has the flexibility to meet swift changes in market demands. These market demands in the view of Shahrokh (1998) are driven by speed of production time. A faster production time is needed due to factors such as delivering the right product to the customer, constant advances in technology that

decrease product life and the high cost of capital. In this agile manufacturing environment the assembly workers need to be trained in various assembly tasks to instil in them a greater understanding of the process in its entirety and this constant retraining is better done on site.

1.5.3 Worker flexibility and training.

Ockerman et al.(1996) saw other reasons for requiring more flexible workforces. Globalisation and computerised production systems resulted in an increase in automation, down-sizing, right-sizing. The increased levels of automation meant that workers had to maintain and keep the automation running. Sophisticated automation systems often required employees to have a specialised knowledge for maintaining these systems. When industries down sized and right sized machinery was brought in to replace human labour. This resulted in fewer workers available to operate the automated equipment. Static technologies such as desktop computers could not provide adequate training and support to the smaller workforce that had to be trained to use sophisticated automated production systems, which may be dispersed throughout a factory site. To solve all the above problems educators and instruction specialists sought to modernise both technological support and training.

In the view of Najjar, Thompson, & Ockerman (1999) this drive for more training in the contemporary work environment meant that training had to become more efficient. According to Gery (1991) traditional training methods had several disadvantages for modern work practices. One problem was that employees had to travel away from the workplace to the training environment. In these training locations the training itself became an “event” performed with images on an overhead

projector representing the actuality of the task or in some cases this reality was represented by a computer programme. These training sessions were led by professional trainers who knew little about the job itself, so often the training session relegated content associated with the task. The presentations used by the trainers became one-size-fits all tutorials adapted to each job with limited application of the knowledge learned about the job to the practise of the job. Experts in the actual jobs were no longer training and mentoring, these specialists were employed in other positions; advising trainers, staffing helpdesks or were promoted into management, away from training altogether.

Carr (1992) comments that training in the early nineties was very costly and time consuming, not immediate and could be largely forgotten by the time that it was required to execute the task. In 1990 American business and industry spent an estimated \$45 billion dollars on formal training programmes. Many companies reported that these training programmes were ineffectual for training basic tasks. Off site training mixed lengthy sequences of background knowledge and overview information and general education with the skills needed for the job. As Royer (1979) points out much of this training was not performed in the context of the job, as a result of this employees had difficulties in transferring what they were learning to their actual job of work. Training was seen to be “training led” rather than providing the knowledge the employee required to do the job. The trainer decided what the learner needed to know, the training was geared towards increasing knowledge rather than improving production techniques and productivity. Traditional training was viewed as being evaluated on learner satisfaction and the achievement of classroom goals instead of increasing job performance. In the view of the above commentators on

training in the workplace good job performance should be the goal of training. Some researchers in the education field like those at the Georgia Institute of Technology envisaged the wearable computer coupled with the appropriate instructional material as a panacea for many or all of the above-perceived drawbacks of traditional training.

1.5.4 Just-in-time training and electronic support systems.

As Kester, Kirschner, Jeroen, Van Merriënboer, & Baumer (2001) explain “Just-in Time” was a production and inventory system that revolutionised Japanese and American manufacturing in the late 1970s and early 1980s. The Fordist style mass production used a traditional “Just-in-Case” (JIC) production system centred on long uninterrupted production runs, stockpiled finished products that had to be replaced by more flexible systems to meet new competitive and economical demands. The Just-in-Time (JIT) is the result of demand-pull production, which uses demand as a trigger for the commencement and the termination of production. This utilisation of demand-pull permits manufacturing industry to produce exactly what is required in the correct quantity and in the right time.

As explained by Globerson & Korman (2001) the continuing climate of downsizing and outsourcing and the use of contingency employees in the workplace have serious ramifications for traditional management policies and work practices. Training strategies is one area that required rethinking and restructuring. The effect of these developments is to reduce worker specialisation of the Fordist assembly line system and to create an increasing demand for workers who can adjust quickly to performing different working tasks that require a wider array of skills. Just-In-Time Training (JIT-T) represents one type of restructuring. Traditional training for certain workplace

tasks may take the form of a seminar lasting five to ten days. Learners would then return to their work environment with the expectation of the company that the training seminar will improve the employees' performances. Not all the new knowledge can be used immediately since only particular issues or procedures can be covered at any one time. A major part of the knowledge gleaned from the training is not used for a considerable length of time. If these skills are not used immediately then they are likely to be forgotten by the time they are needed. The just-in-time training programme is seen as a solution to many of the problems associated with a traditional system of training that has companies sending employees on training courses that provide learners with reservoirs of knowledge to be used when needed. JIT-T provides just enough training and in the right context and constitutes a form of self-paced on-the-job training. However JIT-T still provides a general knowledge element of the subject matter. Table 1.1 compares the assumptions about on-the job training system like JIT-T and traditional methods. JIT-T often is implemented by the provision of an on site study room equipped with workstations.

Najjar et al. (1999) are critical of the way computer-based technologies are used to produce self-paced learning modules. This technique often fails to produce results since the JIT-T ethos is not fully incorporated into the training regime and the instructional developers are using new technology with old training methods. Advanced technology coupled with a new approach in the delivery of learning materials was developed by the Georgia Institute of Technology in the nineties to solve these learning problems in the workplace. This new approach contains all the elements of the learning process including instruction, practice, reference materials and follow up support. This approach is similar to on-the-job training but is an

inexpensive alternative to traditional training and is less problematic to implement for workers who move through a factory site performing tasks. Recent advances in technology have made it possible to create systems that directly aid workers as they train.

Table 1.1: Comparison of traditional versus on-the-job training assumptions. (adapted from Najjar, Thompson & Ockerman(1999)).

Traditional Training Approach	On the Job Training Approach
The information that a worker needs to know is mostly static.	The information a worker knows is very dynamic.
Certain facts, concepts and procedures must be learned before a worker begins to perform a task.	A worker can perform a task poorly at first, but gets better over time.
Training takes place away from the job task, in a classroom.	Training takes place at the job site, while performing the task.
Learning takes place in a specific time period (e.g. the training class).	Learning occurs continually over time.
Learners remember what they learn in class, then apply the knowledge later.	Learners remember what they learn on the job then apply it immediately.
Practice of the actual job does not occur during training.	Practice occurs immediately while performing the job task.
Self directed learning is inefficient and often useless in the workplace.	When given specific goals and rewards for work performance, workers can perform self-directed learning.
After training the worker is an expert at performing the job task.	To become better at the task, the learner continually needs time and guidance from experienced and expensive experts.
There is no evaluation of the effect of training actual task performance.	There is immediate, obvious evaluation of the effect of training on actual job task performance.

The system that was developed by the Georgia institute of technology was an amalgamation of an “Electronic Performance Support System” (E.P.S.S.) and emerging wearable computer technology. As described by Gery (1991) the E.P.S.S. incorporates just-in-time training tasks the learner has to perform and it allows learners to structure their own learning through its interactive features. Its aim is to permit less proficient workers to perform at the same efficiency level as more experienced employees. Modern industry is interested in performance support systems because they purport to improve the performance of the employee and reduce the need for training. The wearable computer training system developed by the Georgia Institute of Technology was called FAST (factory automation support technology). This system has been created to provide employees with the right information or instruction in the right quantity and detail in the right time frame (see table 1.2 for attributes of FAST compared to traditional training).

1.6 Locus of Control.

In systems like FAST, multimedia will be used to deliver the training or instructional materials to the learner. Ockerman, Najjar, & Thompson (1999) describe the advantages of multimedia as providing a range of techniques to present information, including text, graphics, sound, animation and video. Multimedia in their view can provide the user with a personalised perspective of the information allowing the individual to choose the preferred perspective at any given time. Another major perceived advantage is that multimedia can help people learn more information in less time. Media rich interactive learning environments like FAST can be described as having “learner control”. In this environment the learner controls the pace of instruction and the amount of information he or she receives.

Table 1.2: Traditional Training versus FAST (adapted from Ockerman, Najjar, Thompson & Atkinson (1996)).

Traditional Training	Factory Automation Support Technology (FAST).
Training is not integrated with everyday work environment or shop floor process.	Focus on continual learning process in the work environment; not limited to training; assistance provided at moment of need.
Training is done before doing the job task being trained.	Training is done while doing the job task being trained.
Training is focused on increasing knowledge about the job task.	Assistance is provided to improve performance of job task.
Training is trainer-centred; the responsibility for teaching is on the trainer or training system	Learner is responsible for defining learning goals for getting the job done.
Assessment of training is based on learner satisfaction and attainment of classroom objectives.	Assessment of assistance is based on job performance.

Gay (1986) describes computer-based instruction using interactive multimedia as a technology that permits casual access to various video, audio and computer-based information that allows various levels of learner interaction. According to Chapman & Chapman (2000) what makes a computer system with multimedia different from previous forms of combined media is that the elements that represent text, sound, pictures and animations can be regarded by computer programmes as data. This means that the programme can control the order in which the component parts are presented and blended together, and can be dependent on the input of the computer user. The advantage of a learning system that uses multimedia is that the learner can

have more control over the instructions. They may control the pace, depth of study or even manipulate the style of the instruction. This “learner control” computer environment is seen to have advantages for aiding the individual to learn information in their own time and in a way that is more amenable to them. In learner- controlled systems the user can access the materials in any order and may have a choice of multimedia formats to view the learning materials. There may also be advanced features such as connections to the internet or built in subject specific dictionaries and reference materials. This control over instructional materials differs in several ways from traditional learning environments, for example textbooks or videos have an established order in which the information is presented and comprehended. A second type of user control, which is similar to traditional instructional methods, is known as “programme control” This order of presentation is determined by the author or the creator of the instructional programme and is very linear in nature and does not include the advanced features of learner control. This simple linear programme conveys information in way set by the author or programmer. Basic learning materials such as books and videos are seen to have programme control. The use of learner or programme control may be described as the locus of instructional control with the control of instruction categorized as either external (programme control) or internal (learner control).

Researchers who have espoused cognitive psychology have endorsed learner control as an effective aid to learning. Cognitive models of learning have centred on the presentation of information in a way that complements the characteristics of the learner. The proponents of cognitive models also argue that the individualisation of the instructional sequence typical of learner control permits the learner to take great

advances on their own initiative. Other advantages of learner control include the ability to make instructional decisions and experience the effects of these decisions using the instructional materials (Lawless & Brown, 1997).

Despite the existence of substantial empirical evidence that learner control can improve such areas as the performance of learners' attitudes towards the topic and reduce instructional time, there is evidence that in certain circumstances learners perform better in programme controlled environments. According to Hannafin & Sullivan (1996) the nature of the task being learned may have a bearing on the type of control required. Task related situations that mediate the type of control include, whether the task is declarative or procedural, whether the learner has prior knowledge or the task domain and the attitude of the learner to the task. Research has indicated that programme control is superior to learner control for learning procedural tasks. Some commentators, for example, Lawless & Brown (1997) are of the opinion that such learner control environments are not suited to learning procedural tasks especially for novices. However Shyu & Brown (1995) produced evidence from their research that indicated a more complicated picture concerning the situations that favoured learner control. When the learner has prior knowledge and a higher confidence in using the instructional materials they benefit more for a learner control system for any type of task. These more experienced users are better at making decisions about their progress and can monitor their need for more instruction; this ultimately enhances their need for more instruction so ultimately enhances their performance. Gay (1986) reiterated the above postulation by stating that that learner control would be more efficient than programme control only in circumstances where the learners had a thorough comprehension of the content domain.

In an investigation into users attitudes towards interactive multimedia Kettanurak, Ramamurthy, & Haseman (2001) uncovered some surprising facts about users' preferences for the level of control in interactive multimedia, when they tested three specific levels of control. The first level was a non-interactive mode (similar to programme control) where the user moved through the instructional screens in a linear fashion. The second was a low interactivity mode where the user could control the sequence of the of the presentation order. The third was a high interactivity mode in which the user had full control over the programme with feedback from multiple-choice questions about different aspects of the courseware. The non-interactive mode and the high interactivity mode were equally preferred by the participants; however the low interactivity mode was the least liked. Although the high interactive learner control programme was packed with features and options the learners did not find it difficult to use. The simple linear programme control mode was popular since it was easy to navigate and resembled the format of a book or video. The authors of the study speculated that the participants who used the low interactivity mode may have been put off by the fairly inflexible interactive features which may have may have inferred more control that was available.

Comparisons of learner and programme control appear to be have had mixed results in various studies possibly due to the varying operational definitions of learner and programme control. The level of control appears to be a continuum of increasing levels of interactivity and control. Kinzie (1990) cautions that learners will only be able to control their own learning in learner control environments if the learners are properly prepared and if the multimedia instructional systems are well designed. In the view of Lawless & Brown (1997) any decision about the type of control that

should be provided for the learner should be based on the latest models of cognition and research. The courseware that is developed should use computer-based technology within the framework of the current theories. There should be a conscious attempt to tailor the technology to suit the instructional system rather than the instructional system created to suit the computer-based technology. Overall the research reported in this thesis was concerned with “programme control”. The second part of the thesis was primarily interested in the configuration of multi modal information presented on a dynamic visual display, video or animation to enhance learning a procedural task.

1.7 Chapter Summary

- This thesis sets out to discover by experimentation the answer to four research questions in two phases. The first phase of the research programme investigated three research questions; the first was whether using head-mounted displays for following instructions is the same as traditional instruction formats. The second question examined which is the best subtype of head-mounted display for following instructions for executing a procedural task. The third research question compared the efficiency of see through head-mounted display screens to opaque screens for following procedural instructions.
- Difficulties in comparing head-mounted display subtypes in the first experiments led a second phase to concentrate on the optimum presentation of media in a multimedia instructional programme for learning procedural tasks using computer-based technology. One reason behind this research is to

ascertain whether the guidelines for learning declarative knowledge are the same for learning procedural tasks. This is accomplished by testing through experimentation whether the optimum configuration of media in a multimedia programme for learning procedural information is the same as for declarative knowledge. A second reason is conditional on the first; if there are differences in these media configurations, these differences could be used as guidelines for the optimum configuration of media for learning a procedural task.

- The background to the research involves changes in the type of production in industry in recent years. These changes have impacted upon training in the workplace. Traditional methods of training in the classroom environment had become costly and inefficient for repeatedly training workers in new tasks. Training has to become more cost effective and specialized, manufacturers have responded to this by using emerging technologies such as the wearable computer to make training more efficient.
- Multimedia has become an important facet of computer-based training and emerging technologies. There are two major levels of control in multimedia computer-based instructional packages; learner control and programme control. The former allows the user more interactivity with the package and organise the pace and structure of their learning. The latter is more linear in construction and is restricted to a set order and tempo of instruction. Learning procedural tasks may be better-learned using programme control.

Chapter 2

2 Head-mounted display technology.

“Let's imagine a new approach to computing in which the apparatus is always ready for use because it is worn like clothing. The computer screen, which also serves as a viewfinder, is visible at all times and performs multimodal computing (text and images). “(Mann 1997, p 7).

2.1 Emerging technologies for industrial training.

Phase one of this thesis examined one particular emerging technology that is used for training, the wearable computer. This chapter will describe wearable computer technology and the issues involved with its usability for performing tasks. The chapter will then discuss research carried on head-mounted displays. A third section will review research on head-up display technology. Lastly there will be a description of the head-mounted display technology used in the first phase of experiments. The research team at the Georgia Institute of Technology had a wearable computer system custom built to access their electronic performance support system. Their wearable system was typical of mobile systems developed in the mid to late nineties. Thompson, Najjar, & Ockerman (1997) describe this wearable system and the rationale behind the research they carried out at the Georgia institute. The wearable computer system consisted of two major components, a belt-mounted microcomputer and a head -mounted display.

Figure 2.1: Principle components of wearable computer developed by Georgia institute of technology. (Adapted from Thompson, Najjar & Ockermann (1997)).

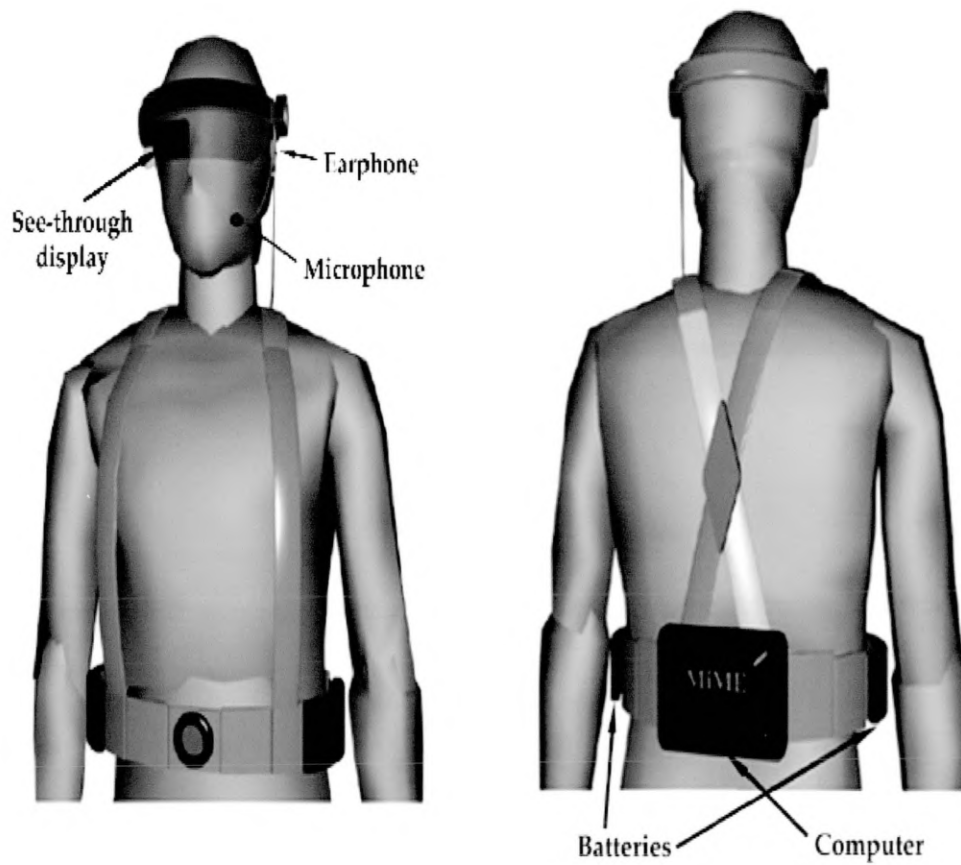
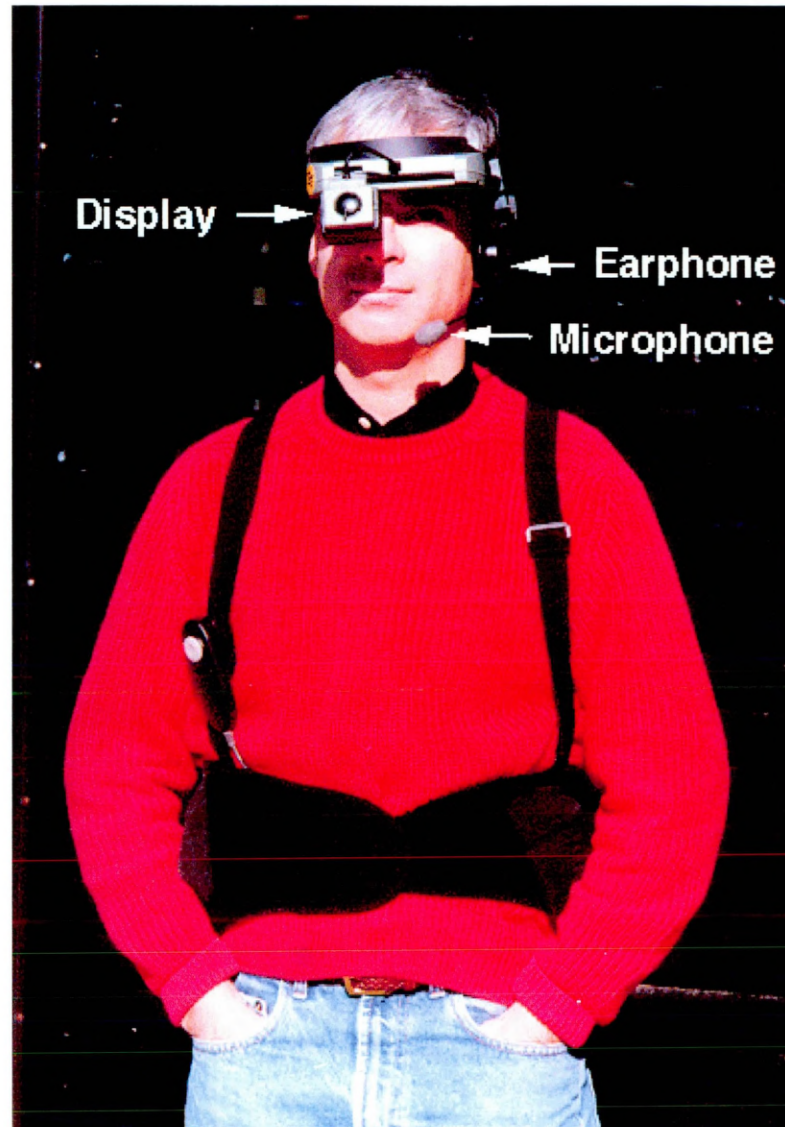


Figure 2.1 shows the front view and back view of the main components of this wearable system. This belt-worn computer was comparable to a desk top unit and could run any operating system run on a P.C., e.g. Window, Unix, or DOS. Information was entered into or retrieved from the computer using either voice control, a hand held pointing device or a wrist-mounted keyboard. As with other wearable systems the head-mounted display operated as the computer screen. Figure 2.2 shows the prototype, which used a monocular head-mounted display. The wearable computer coupled with a multimedia learning system such as EPSS was heralded by many commentators as a future facilitator for training skills in the workplace.

2.2 Head-mounted display technology.

The developers of wearable computer systems are concerned that the average worker can be trained just as efficiently with these wearable systems compared to traditional training methods such as paper-based instructions and desk-top computers. If wearable computers systems prove to be ergonomically difficult for users then this would negate any advantages these systems would have over more conventional training materials. Barber, Haniff, Knight, & Cooper (1999) identify the head-mounted display as a central component of the wearable computer system and state that cautious deliberation is required for the design of head-mounted displays and the information that is displayed on them. Melzer & Moffit (1997) define this technology as an image source with collimating optics in a display mounted on a user's head. The head-mounted display may have either one or two display channels that project graphics and symbols. The information may be viewed directly in an enclosed unit

Figure 2.2: Prototype wearable computer developed by Georgia institute of technology.(Adapted from Thopmson, Najjar & Ockermann (1997)).



that excludes an external view of the outside world for an immersed experience or they may take up only a part of the wearer's visual field and allow a partial view of the outside world. These head-mounted displays may also use a semitransparent or transparent display that permits the user to look through the display to the outside world. The head-mounted display could be a monocular system with the image projected to one eye or the image may be projected to both eyes as in binocular or biocular systems.

Ellis & Menges (1998) report that most of the literature concerning the design of display technology has been concerned with head up displays (HUDs) in military aviation. These head up displays project images that focus at more than two meters from the users eyes so that the display image is superimposed on the outside environment exterior to the cockpit window. By the early 1980s the head up display had become established as a means of presenting both alphanumeric and pictorial flight systems to pilots in military aircraft (Gibson, 1980). As illustrated by Roscoe (1993) by the early 1990's head-mounted or helmet-mounted displays began to be used in military cockpits. These headsets serve the same purpose as the head-up display and project aircraft symbols such as pitch ladder and artificial horizon on to the real world. However with a head-mounted display the projected information moves in the user's visual field with head movements. The image source in the head-mounted display may be a cathode ray tube or a liquid crystal display. The latter display type is becoming more prevalent in head-mounted displays due to its small size; this makes the headset smaller and weigh less. Liquid crystal display (LCD) technology has been used for many years for displays in digital watches and small hand-held television sets.

As Travis (1991) points out the advantages of this miniature display technology is that it provides high spatial resolution and the displays are thin and lightweight being only a few centimetres thick. These liquid crystal displays operate with a low power so are efficient with energy. The recent appearance of modestly priced head-mounted displays has extended the gamut of potential applications of display technology. These potential applications include learning mechanical assembly, surgical preparation, surgical training and the visualisation of objects in computer-aided design. The emergence of these new head-mount displays has introduced issues about optical design and information presentation that have not previously been dealt with.

According to Melzer & Moffit (1997) one of the dilemmas in the design of head-mounted displays is that apart from some rough guidelines there is a lack of specifications that set the parameters of human performance using this technology. There is no real body of human factors research to draw upon when designing a head-mounted display. From an ergonomics point of view, the user should be considered an integral part of the entire head-mounted display system. Specifically, the design should take account of the human visual system where the eyes act in tandem as a binocular pair on a moving head. Human vision is limited by the nose and the forehead, creating a central binocular visual field with flanking monocular visual fields. For humans the eyes move on the head that rotates on the body that in turn moves in space, this head movement necessitates particular system requirements for the design of a head-mounted display. There are four main areas of ergonomic design pertinent to head-mounted displays; the first is the upper limits of various visual requirements such as focus, luminance, alignment, colour and resolution. The second is the physical requirements of each user of the head-mounted display, since

individuals have different shapes and sizes of heads and the design must take account of this. The third issue is an environmental one that deals with such design issues as the comfort of the user in terms of fit and temperature in the head-mounted display.

The fourth main issue is the accessibility and ease of use of the controls.

As pointed out by Lin & Kreifeldt (2001) the lack of ergonomic design considerations extends to the entire wearable computer system and not just the head-mounted display. Technological advances have given companies the opportunity to develop and manufacture wearable computers. However ergonomic issues concerning wearable computers are being superseded by technical considerations. These wearable computers may be uncomfortable to wear or interfere with normal physical activities. The majority of head-mounted displays incorporated into wearable systems are monocular. Rohaly & Karsh (1999) recount that the U.S. military pioneered several research programmes to develop the head-mounted display for use in aircraft. The results of this research produced several reasons for using monocular rather than binocular head-mounted displays. These reasons include reduced cost and weight, the assumption that there will be an increased field of view, a decrease in spatial disorientation and simplify navigation. As well as the above arguments there is the assumption that the monocular head-mounted design will decrease workload by allowing two tasks to be performed simultaneously by allowing each eye to perform a separate task. However few studies have been conducted to explore dual-task performance under such dichoptic viewing conditions since this is not a naturally occurring situation in the human visual system.

2.2.1 Cognitive Issues with head-mounted displays and head-up displays.

The first phase of experiments reported in this thesis was concerned with investigating the use of head-mounted-displays for following instructions to execute an assembly task. This first phase explored a number of research questions; these questions were all concerned with the participant's eye and head movements whilst switching between viewing the instructions and executing the task. The first phase investigated the properties of two different types of head-mounted display and a head mounted display in opaque and see-through mode. The initial research question posed in the first phase was whether head-mounted displays are more efficient for following instructions than traditional methods. A possible advantage of head-mounted displays is that they may lessen the amount and length of eye and head movements between task and visually presented instructions. This issue of economical eye and head movements was a major reason behind the development of head-up and head-mounted displays in military aircraft. Foyle, McCann, Sanford, & Schwirzke (1993) explain that head-up displays in cockpits were developed to permit pilots to retain awareness of both cockpit controls and the outside environment. The head-up display uses the method of superimposing aircraft symbology at optical infinity in the pilot's field of view. This allows pilots to access both the outside world and primary onboard aircraft displays in the same region of fixation and accommodation.

Many studies, for example Weintraub, Haines, & Randle (1993) demonstrated that pilots have less head and eye movements using head-up displays compared to those who used conventional head-down displays in cockpits. The concept behind head-up display development is that by decreasing head and eye movements and increasing the

time spent looking at the workspace, there will ultimately be an enhancement in the performance of tasks. The overlaying or placing of essential information on the workspace in a spacially meaningful way reduces time taken searching for specific details. One question addressed in this thesis is whether head-mounted displays with opaque or see through screens can demonstrate a similar reduction in attention switching between the instructions and the task. Some theorists have suggested a strong association between spatial location and short-term or working memory. Kirsh (1995) argues that techniques that manage space are fundamental for organising human cognition and behaviour. When space is used properly, time and memory demands for tasks can be brought down to manageable levels. The reliability of the performance can be augmented as well as the number of jobs performed simultaneously. By spacially arranging information together with task environment, the head-up display or head-mounted display may furnish a strong control of spatial cognition and memory (Ververs & Wickens, 1998).

However there is an argument that the method humans use to coordinate movements of both the eyes and the head to scan or search the visual field may pose problems for using head-mounted displays to view information. According to Uemura, Arai, & Shimazaki (1980) when the scanning distance to a point only requires a small angular movement, the eyes move first followed by the head. When the angular distance is larger the head usually moves in conjunction with the eyes. Leigh & Zee (1999) contend that the inclination to execute an eye-head movement depends in part on the ocular motor range, which in humans is approximately plus or minus 50 degrees from the centre of the visual field. Targets that are scanned outside this range generally need eye-head movements to reach them. There are however individual differences in

the tendency to make these eye-head movements to visually acquire objects at a particular angle. This mechanism of coordinated eye-head movements may cause problems whilst using head-mounted displays since these units are attached to the user's head. When viewing information on a head-mounted display, head movements will not centre the display in the visual field; all scanning must therefore be done with the eyes. Attempting to read information with the eyes continually off centre may cause eyestrain. Peli (2001) indicated that this latter characteristic of the head-mounted display may create problems for scanning material placed close to the edge of the screen. As a result of this viewing information with visual angles greater than 10 degrees from the centre would be uncomfortable to sustain.

2.2.2 Previous Research with head-mounted displays.

Wearable computers constitute a nascent technology and there is a dearth of reported studies in the research literature that test the usability of the head-mounted display component of a wearable computer Barber et al.(1999) compared a head-mounted display to a desktop computer for following instructions to accomplish a task. The experimental task was to solve the puzzle game Solitaire. Participants followed on screen instructions that helped them solve the puzzle on an actual game board. Both display conditions required that participants divide their attention between the game board and the display. The results of this study indicated an overall slower performance for the head-mounted display in comparison to the desktop computer. It was concluded that in the head-mounted display condition participants took longer and required more effort to retrieve information from the monocular display. The study used a monocular head-mounted display; this has different properties to a binocular head-mounted display. Only one eye can view the display in a monocular

display system and the other eye is stimulated by the real environment. This means that eye movements from the task to the display are different to a binocular system. A monocular system has a more limited display space; this reduces the amount of information that can be presented on the screen. A monocular display is also susceptible to binocular rivalry. This occurs when the each eye is stimulated with a dissimilar image to the other and this causes the two different images to alternate backwards and forwards in the visual field. (Blake, 1997).

Rohaly & Karsh (1999) tested the efficiency of an opaque monocular head-mounted display for a dual target location task. They investigated the claim by head-mounted display manufacturers that a monocular system is more efficient for dual task performance. Their findings indicated that having participants perform a visual task using one eye and perform a second task with the other does not improve performance on these tasks and is less effective than having one eye do the tasks. The eyes do not function as independent channels so a monocular system divides attention and decreases dual task performance. As described by Gregory (1990) the human visual system is complex and starts with the retinas of the eyes, which are divided vertically into two parts. The optic nerve fibres from the each inner part of the retinas cross at an area of the brain known as the optic chiasm on their way via the thalamus to the striate cortex. Fibres from the outer parts of the retinas on the other hand do not cross. As explained by Goldstein (1999) this system of connections allows visual representations of an object to be processed in the same area of the cortex.

Only a few studies have compared a wearable computer system with a conventional vehicle for presenting instructions to execute a procedural task. Ockerman, Najjar, &

Thompson (1997b) compared an electronic performance support system on a wearable computer with a monocular see-through head-mounted display to a book for presenting information on origami and instructions for folding paper to create the origami shapes. The results from their study revealed that participants in the wearable computer condition took significantly longer to execute the task but made fewer errors than participants who used the book instructions. The book users were able to look ahead in the instructions and generally move backwards and forwards quicker whereas the wearable computer users can only move a step at a time.

Users had problems with some of the elements of the interactive system on the wearable computer due to being unfamiliar with the complex interface design. The electronic performance support system installed in the wearable computer had an interactive structure with a locus of user control that was based on a learner control format (see section 1.6 for a description of locus of control). This included various utilities such as a self-correction programme, a choice between step-by-step still images of the procedure or a video clip showing the shapes being folded by hand. It also included a dictionary of origami terms and features. This type of learner control programme may have proved too complicated for a novice user and less efficient for following procedural instructions (Lawless & Brown, 1997). The head-mounted display used in this study was a see-through or transparent monocular display. The eye and head movements required by this type of display may have been one of the factors that affected the performance of the participants in the experimental task. This observation is based on one of the findings of the (Barber et al., 1999) study that reported that participants were slower switching their focus from the display to the task on a monocular head-mounted display (See section 1.3.2 for discussion of eye-

head movements). The first phase of experiments tested whether binocular displays produced similar results to monocular displays and demonstrated some inefficiencies in processing displayed information.

Very few studies have compared the efficiency of opaque and see through screens in head-mounted for executing tasks. Laramie & Ware (2002) demonstrated through employing a simple reading task that see through monocular displays were unsuitable when compared to opaque monocular displays for use in busy dynamic environments. See through monocular displays require a static background and are prone to perceptual problems in busy or moving environments. For example see-through monocular displays are prone to visual interference, a situation when two images are not easily distinguishable from each other and the user is unable to distinguish them visually.

2.2.3 Previous research with head-up displays.

Although empirical research into head-mounted displays is sparse, there is a body of work in the literature that has investigated the use of head-up displays in aircraft and motor vehicles. In one particular study, Sanford, Foyle, Mccann, & Jordan, (1993) identified a problem with head-up displays that occurs when users have to process head-up display symbology and the outside world concurrently. The inability to process these two pieces of visual information efficiently was caused by increased eye-movements needed to scan different elements of the head-up display symbology, this caused a divided attention effect between elements of symbology and the outside world. Another problem with head-up displays that has been noted is that pilots do not

always attend to both the head-up display symbology and the outside environment simultaneously.

Fischer, Haines, & Price, (1993) demonstrated that pilots using the head-up display to land aircraft consistently failed to notice obstacles on the runway, suggesting a failure to monitor both the symbology and outside environment efficiently. Although head-up displays have on occasion been shown to improve performance, some studies have demonstrated that feedback from participants is often negative towards head-up displays. Sojourner & Antin (1990) compared the effects of a simulated head-up displays on driving tasks in a motor vehicle simulator to a condition without the head-up display. Despite superior performances in experimental tasks in the head-up display condition, participants reported that they did not find the head-up display useful for monitoring navigation and speed. Studies from the selective attention literature have questioned the efficiency of looking between two information sources when one source is superimposed on another. Becklen & Cervone (1983) had participants attend to only one of two visually superimposed video taped baseball games. An unexpected event went unnoticed in the unattended ball game by a majority of the participants. This finding indicates that optically superimposed information may be difficult to process in parallel.

The experiments undertaken in phase one of this research thesis set out to build on previous research in head-up and head-mounted displays. Specifically these experiments were designed to address the first three research questions this thesis set out to answer. These experiments set out to test for differences between traditional instructional materials and head-mounted displays, different display types and two

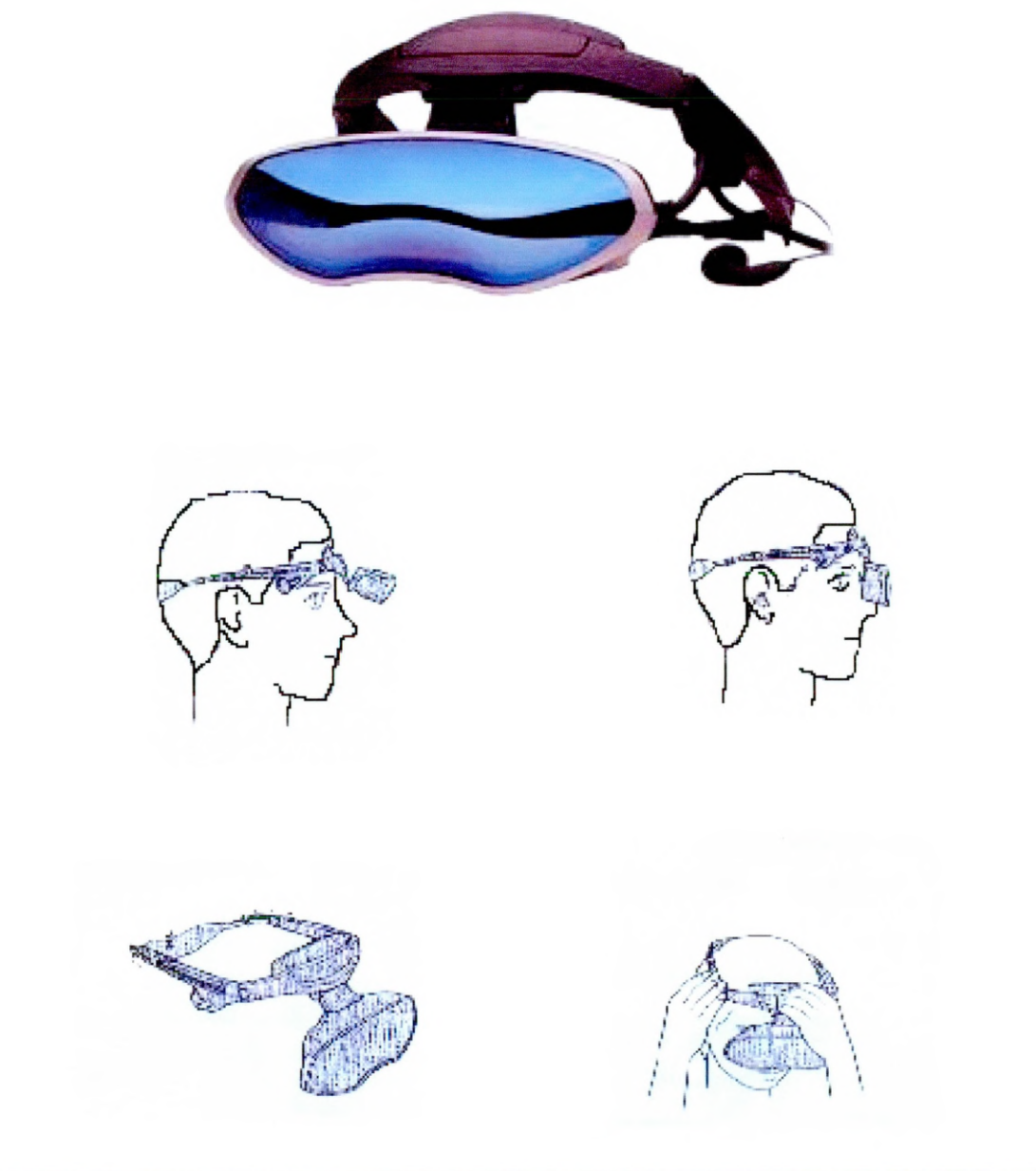
different screen types. The task performed in these experiments involved assembly instructions for constructing abstract Lego models. This task was chosen since it typifies a procedural task encountered in the work environment.

2.3 Head-mounted display technology used in the first phase.

A Sony Glasstron head-mounted display (See Figure 2.3) was used to investigate the first research question, the comparison of the head-mounted display to traditional vehicles of instruction. The Glasstron is equipped with two 1.55 million pixel LCDs (Liquid Crystal Displays) and can be connected to a P.C. The screen of the P.C. can be seen on the liquid crystal display in SVGA mode with a resolution of 800 by 600 only. The field of view of the Glasstron is about 30° and the image is projected at approximately five feet. The LCD display screens are attached to the headband by a hinge so that they can be lifted from the eyes. The Glasstron can be worn with glasses, however there is no option for adjusting the inter-screen distance on the Glasstron.

The Glasstron is a binocular display that has two display modes; opaque and transparent. To answer the first research question the Glasstron was used with its screen in opaque display mode. Whilst wearing the Glasstron users can switch their gaze from the display to underneath the head-mounted display where their remaining visual field is large enough to permit them to monitor themselves performing a task. Text information was used in the animation in this experiment to describe the dimensions of the bricks in order to have the same multimedia instructions in each condition. The text and picture format of the paper instructions were recreated in the animation used in the computer-based conditions. The animation featured the same pictures of the stages of the construction .

Figure 2.3: Sony Glasstron Head-mounted display. (adapted from Sony instructions for the LDI-100BE 1999).

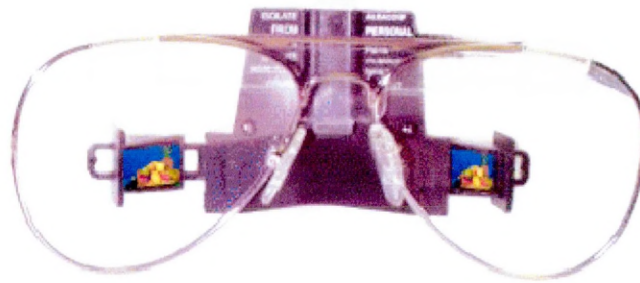


of the Lego models as the paper booklet, so that the instructional materials in the three conditions were more equivalent.

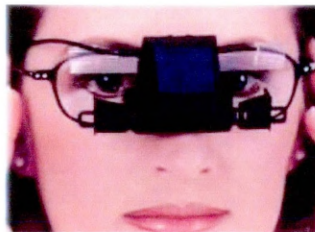
The second question addressed by this experimental programme concerned testing the efficiency of two head-mounted displays with different display positions. The same head-mounted display, a Sony Glasstron that was used in the previous experiment was tested against a second head-mounted display, the Albatech Personal Monitor. The Albatech is a binocular head-mounted display that produces a “floating television image” six feet in front of the eyes with a field of view of about 20 degrees. The image can come from any equipment: T.V. monitor, camcorder, video-recorder or a P.C. video card that can generate a composite video signal PAL or NTSC. At the centre of this system is one liquid crystal display. The image of the display is conveyed by an optical system into both eyes of the viewer. Despite the fact that the images are arriving into both eyes separately the optical system allows the brain to merge the two images into one. This is achieved for each individual user by adjusting the two mirrors until the user can see one image (See Figure 2.4 A, B and C). The display clips on to a pair of eyeglasses. If the user does not require glasses prescribed from an optician, an eyeglass frame is provided (See figure 2.4 D, E, F and G).

Not only do the two displays require different eye movements but they also project images with dissimilar fields of view. When using the Glasstron the user must move their eyes from an opaque display screen that occupies most of the top part of their visual field down to the task area that takes up the bottom part of their visual field. The Albatech head-mounted display projects an opaque image with a much smaller field of view directly in the middle of the user’s visual field. To view the task area the

Figure 2.4: The Albatech Personal Monitor
(Adapted from PM-1B-CFLI-PAL/NTSC user manual 1997)



A



B

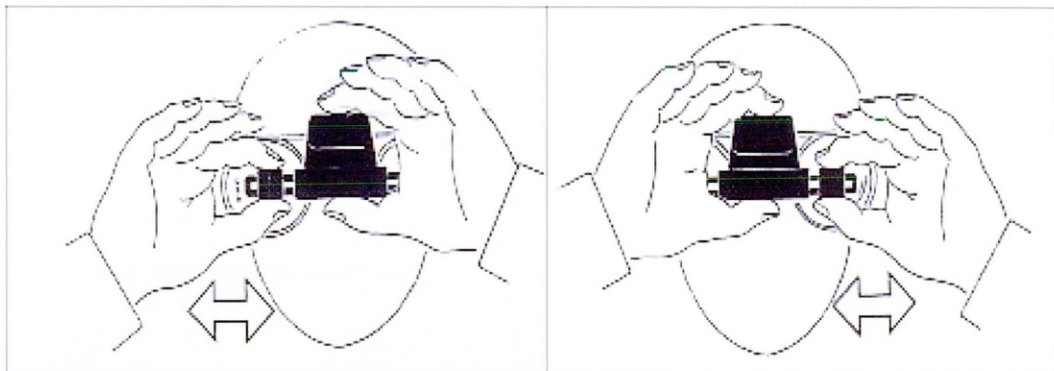
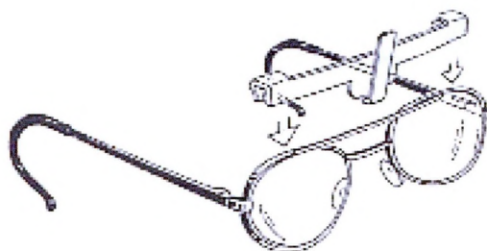


Figure 2.4: The Personal Monitor PM-1B-CFLI-PAL/NTSC (continued).

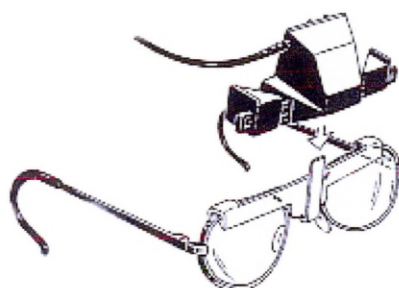
C



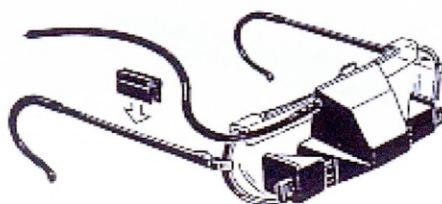
D



E



F



G

user must glance past the projected image to the task area. In this experiment a video clip showing a pair of hands in close up building the Lego model replaced the animation used in the previous experiment. The video clip was used since it gives a clearer depiction of the parts needed for the assembly and conveys more action information necessary for a procedural task (Park & Hopkins, 1993). The video was played on Windows media player that has controls that resemble those on a videocassette recorder, this interface would be more familiar to participants and more intuitive to use.

The third question explored the possible differences in efficiency for following assembly instructions on an opaque display and following the same instructions on a see through display. See through displays are ubiquitous in wearable computer technology and are used in both monocular and binocular systems. This experiment set out to test if the different head or eye movements involved in using see-through and opaque systems would have an affect on task performance. In this experiment the Glasstron was used for both conditions since this head-mounted display has an opaque mode and a see through mode. Given the possible perceptual problems of reading text information on the see through screen, the complementary information for the video clip in the third experiment was presented as a soundtrack of a voice.

2.4 Chapter Summary.

- Wearable computer technology was the emerging technology tested in the first phase of experiments. The experiments in phase one were planned to focus on the head-mounted display component of the wearable computer investigating some of the ergonomic issues concerning this developing technology. There are several visual problems with viewing information on the liquid crystal displays inside the head-mounted displays that need to be addressed.
- There are possible advantages of using head-mounted displays for viewing training information. The main advantages of using a head-mounted display are the economical eye and head movements for viewing information whilst executing a task. Disadvantages of using a head-mounted display include problems with head and eye coordination whilst reading and viewing information on the display screen. Eyestrain may be caused by trying to read information on the edges of the screen that may become distorted.
- There were a few studies into using head-mounted displays for reading information and executing tasks. These studies compared head-mounted displays to traditional vehicles of information. A majority of these studies tested monocular headsets. Findings from these studies highlighted visual and ergonomic problems with monocular head-mounted displays. There were also studies that explored the effectiveness of head-up displays for performing tasks in aircraft and motor vehicles. Some of these studies also highlighted visual problems including split attention effects.

Chapter 3

3 An Investigation into using head-mounted displays for following multimedia procedural instructions for an assembly task.

“Most of the greatest advances in modern technology have been instruments which extended the scope of our sense organs, our brains or our limbs.”

(Craik,1952, p.61)

3.1 Phase 1 of experiments.

This chapter reports the first phase of experiments undertaken in this research. These experiments investigated features of binocular head-mounted displays, specifically regarding their physical use for following instructions. In this chapter there will be a description of each experiment in the first phase and the results from each experiment will be reported. The rationale behind the first phase of experiments centred on three research issues. The first issue investigated whether the usability of wearable systems compares favourably with traditional training methods, especially for the presentation of information on the head-mounted display component. The second issue was concerned with ascertaining which head-mounted display design is best suited to training tasks in the workplace. A third important issue is which type of display screen is the best for following instructions in a head-mounted display; opaque or see-through. At the end of the chapter there will be a summary of results of phase one that examines what implications the findings from the experiments have for the research issues.

3.1.1 Programme of experiments in phase 1.

Table 3.1: Programme of experiments for phase 1 of experiments.

Experiment	Conditions in experiment.	Rationale for experiment.
1.1	Instructions on Opaque Glasstron, desktop P.C. and paper.	Compare effectiveness of head-mounted display to traditional vehicles of presentation.
1.2	Instructions on Opaque Glasstron, Albatech and desktop P.C.	Compare effectiveness between two types of head-mounted display with different display positions and desktop P.C.
1.3	Instructions on Opaque Glasstron, Transparent Glasstron and desktop P.C.	Compare effectiveness between an opaque head-mounted display, transparent head-mounted display and desktop P.C.

Table 3.1 outlines the conditions in each experiment in phase 1 and the rationale behind each experiment. This series of experiments investigated the efficiency of the head-mounted display component of a wearable computer for following animated instructions to build abstract Lego models. Each of the experiments in phase 1 set out to investigate a specific research question.

Experiment 1.1 compared the effectiveness of the binocular head-mounted display component of a wearable computer system to more conventional methods of delivering assembly instructions to build an abstract Lego model. A desktop computer and a paper booklet were the conventional methods that the head-mounted display was measured against. Experiment 1.2 set out to determine whether the position of the screen in the head-mounted display had an affect on following the instructions. The position of the screen may have an effect on the eye movement involved between reading the instructions on the screen and monitoring the task being performed. A screen placed more centrally in the user's field of view may be more efficient for

following instructions since this may decrease eye movements, as was the case with HUD technology (see section 2.2.1). The final experiment in this phase, experiment 1.3 investigated whether there is a difference in following the instructions when the screen in the head-mounted display is opaque or see-through. It is important to ascertain whether looking through the instructional video played on the screen in the head-mounted display will help or hinder the task.

These first experiments used a simple linear animation with a locus of control based on a programme control system. The rationale for this experimental paradigm was to evaluate the efficiency of binocular head-mounted displays for presenting multimedia instructions. The dependent variables in these experiments were completion time and error rate. The first was a measure of how well the participants could take in the multimedia instructions in the various experimental conditions and keep up with these animated instructions. The second was a measure of the accuracy of task performance in the different conditions in the experiments.

A different set of Lego models were used in each of the three experiments due to the difficulty in acquiring large enough participant numbers for the experiments. Some of the participants took part in one or more of the experiments. If the same three abstract Lego models had been used for all three experiments practice effects would confound findings from the experiments. The participants who took part in more than one experiment would have become practised at building the three Lego models.

3.2 Experiment 1.1: Efficacy of HMD Compared to Desktop and Paper for Presenting Instructions for Assembling an Abstract Model.

A pilot study using six participants, three males and three females, set out primarily to test the animated instructions that were created for the first experiment. The visual representations of the abstract models in the pilot were positioned at different distances and angles to one another in the animation and still frame components of the interface as well as the instruction booklet. The pictorial instructions were designed so that Model A should be easier since it was at a closer distance and the view of the model was almost a straight side elevation. Model B should have been slightly harder since the view was at an intermediate distance and at an angle that may have made it harder to determine the orientation of the bricks when placing them on the model. It was hypothesised that Model C would be the hardest to build since the view of the model was more distant and seen at an angle (see appendix 3.1).

Lego was used to construct the models in the pilot as well as the other three experiments in phase one since Lego allows the construction of many abstract models with a similar level of difficulty. In one condition in the pilot the instructions were simple paper instructions (see appendix 3.2). The other two conditions had the paper instructions converted into an interactive computer programme. The instructions in the pilot study that were computer-based were in the form of a basic screen interface with two windows, one with an animation sequence showing each brick and where it was to be placed on the model, alongside a second window with a step by step still frame of each stage in the procedure of the construction of the model (see appendix 3.3). This latter feature was to give participants the option of using the still frame window like a book and move through the stages backwards and forwards.

The pilot study had a qualitative section that involved a short semi-structured interview with each participant after they had completed the experiment (see appendix 3.4). The participants were asked six specific questions about the mechanics of the experiment. The seventh question invited general comments about the experiment from participants. Due to the small number of participants statistical analysis was inappropriate. The results of the semi-structured interview can be seen in appendix 3.5. An important finding from the pilot study was that no participants used the frame-by-frame feature but just followed the animation whilst building the model; the frame-by-frame feature was dropped from the animation in experiment 1.1. The links to the stages were placed down the side rather than the bottom this was another change to the animation that came from participant feedback, this would make the interface easier to use. Two participants reported difficulties following the instructions on the head-mounted display compared to the desktop. They found that the instructions were either harder to see clearly or smaller than the desktop. One participant also reported having difficulty discerning the orientation of the bricks in the head-mounted display condition. These participants may have sat nearer the desktop monitor giving them a closer view of the animation compared to the image on the head-mounted display. To counteract this, participants in experiment 1.1 sat a standard distance from the desktop whilst viewing the animations.

Some participants said that they found the head-mounted display awkward and difficult to use at first but they eventually got used to it. The fact that most showed a preference for the paper instructions may indicate that this is a format for instructions that they are more familiar with. These findings influenced the way experiment 1.1

was set up. The pilot was mainly concerned with the usability of the interface for viewing the animation. For experiment 1.1 video cameras were used to record the performances of the participants to allow a full analysis of their performance and measure not only the time taken to perform the task in each condition but also to count the number of errors made in each condition.

Experiment 1.1 extended the pilot study by including a quantitative section that investigated whether there were any differences between the efficiency and usability of a wearable computer compared to traditional methods. Twenty-one participants were asked to build three abstract models from Lego bricks in the same three conditions as the pilot, with the instructions on a desktop, on a head-mounted display or on paper. The three abstract models used in this study were of a similar construction and had the same number of bricks. The pictorial representations of the three models in the instructions had the same angle and size as the model in the pilot study that participants found the easiest to build. After problems in the pilot study with possible inequalities in viewing distances between the head-mounted display and the desktop P.C., participants viewed the interactive instruction programme in the desktop condition with their heads fifty-four centimetres from the computer screen. This standard distance from the screen gave them the same field of view as in the head-mounted display condition.

The participants were filmed building the three models and their completion times were calculated during video playback and analysis. The use of video equipment permitted the introduction of a second dependent variable the number of errors they made performing this task, this would be a measure of accuracy in performing the

task. After completing all the conditions in the experiment the participants were asked about their views on the experiment using a semi-structured interview.

3.2.1 Method.

Design

This study was in two parts; it had a quantitative section and a qualitative one. The quantitative section employed a within subjects design. The independent variable was the instruction format used to build the models; paper, head-mounted display or desktop P.C. The dependent variables were a) the time taken to construct the models in seconds and b) the errors made whilst building the model. In this study the criterion for an error was the placing of a brick on a model in the wrong place or in the wrong orientation. If a participant placed a brick wrongly but realized their mistake and corrected their error before the model was completed, this was not counted as an error. The qualitative section involved a semi-structured interview with six set questions and one open ended question to obtain feedback from the participants about the experiment.

Participants.

There were twenty-one participants; eleven males and ten females with an age range from eighteen to forty-five. The participants were students of the University of Abertay Dundee. All the participants were all unpaid volunteers. All the participants had normal or corrected eyesight. Originally thirty participants were tested. However the data from nine were omitted from analysis because they did not wear their eye correction in the head-mounted display condition. These nine participants did not normally wear eye correction for viewing a computer screen so neglected to wear eye

correction for the head-mounted display condition. This only became apparent after the experiment and was disclosed at the post-test interview. The performance of these participants may have been compromised since they may have needed their glasses for distance. This group was following the instructions on a projected image at two meters and therefore may not have seen the image clearly.

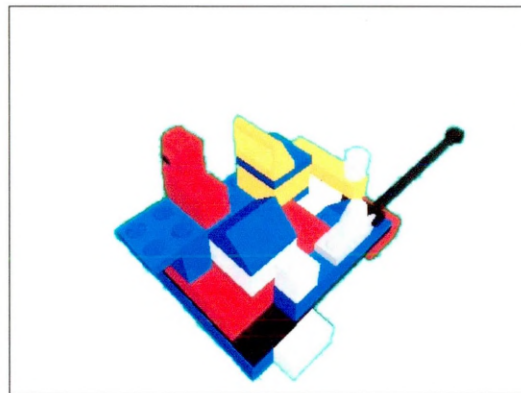
Materials and Apparatus.

Materials in this study comprised of two computers, a Sony Glasstron binocular LCD head-mounted display that could view output from PCs. Play bricks to construct three models (See Figure 3.1). The models in the study were referred to as “Alpha”, “Beta” and “Gamma”. Each model had 27 bricks and had similar characteristics in construction. There was a similar angle of view of each model in the pictorial instructions, type of bricks and complementary text instructions. These similar features in three models it was proposed would give them a similar level of difficulty in following the instructions to build the models. The paper instructions were similar to the pilot study and consisted of a booklet with the twenty-seven pictures of the stages required to build the model. An interactive programme using the same instructions as the paper guide was used in the desktop and head-mounted display conditions (See Figure 3.2). This comprised of a screen presentation, with an animation of the model on the right hand side and links to stages on the left. Participants could go back to any stage in the animation if they fell behind whilst building the model. A camera was used to film the participant’s hands whilst constructing the models (See Figure 3.3B).

Figure 3.1: Abstract models used in Experiment 1.1



Alpha



Beta



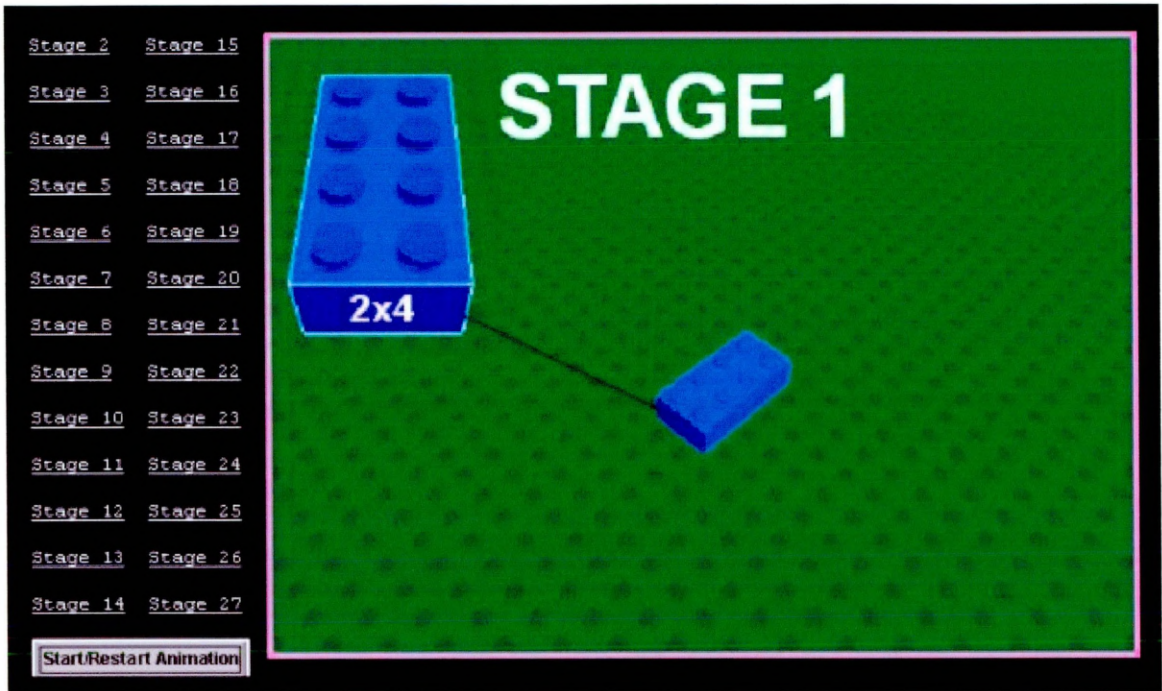
Gamma

Procedure.

The order in which participants were assigned to the conditions was counterbalanced (see Appendix 3.6). The three types of display in which the instructions were presented to the participant were paper, desktop and head-mounted display. The same instruction programme was used in both the desktop and head-mounted display conditions. One PC was used for the desktop condition and the other for the head-mounted display condition.

Before using the desktop and head-mounted display instructions, the experimenter carefully explained the control system to the participant. The participant was asked to sit in one of two testing bays (See Figure 3.3 A). One bay was used for the head-mounted display condition, the other testing bay for both the paper and the desktop conditions. In the latter condition the participant's head was positioned fifty-five centimetres from the screen to give participants the same field of view of thirty degrees as they would encounter in the head-mounted display condition.

Figure 3. 2: The interactive programme for Desktop and HMD conditions in Experiment 1.1



The instructions comprised of a set of unpublished web pages created using Liquid FX v 3.1 HTML editor. The stages of construction of the models were made using Lego Creator. The model was constructed in the program. Then each stage was saved as a jpeg file using the screen capture facility of Paint Shop Pro v.6. The jpeg files were then turned into animations. The animation was made using Animation Shop v.2. that ran at a speed of one frame every 4.5 seconds. The right hand panel contained the instructions in a frame-by-frame format. The left-hand panel contained links to each stage in the animation. The current piece to be used is in the left-hand corner of each panel, with an arrow indicating its position on the model. The "Start/Restart Animation" button starts the animation from the beginning.

Figure 3.3: Testing bay with cameras used in experiment 1.1 Efficacy of HMD Compared to Desktop and Paper for Presenting Instructions for Assembling an Abstract Model.

A



In experiment 1.1 only one camera was used to record each subject's hands close up as below.

B



Still From video clip of participant in experiment 1.1

In the head-mounted display and computer conditions the participants were asked to wear their eye correction, if they normally did so to look at a computer screen. The participants were asked to construct a different model out of play bricks in the three conditions this was done to counteract practice effects. The participants were also asked to build the model in a predefined area to ensure that their hands would be captured by the video camera. After the participant had completed the test battery, the experimenter asked a set of questions about the experiment and noted the responses on a feedback questionnaire similar to the pilot study. The questionnaire took a semi-structured format and comprised of seven questions to elicit feedback about the mechanics of the programme and about the three instruction formats (see Appendix 3.7). The completion times and errors were calculated during video play back.

3.2.2 Quantitative Results.

Figure 3.4: Graph showing mean completion times and standard error in seconds for each instruction condition: Paper, Desktop and Head-Mounted Display.

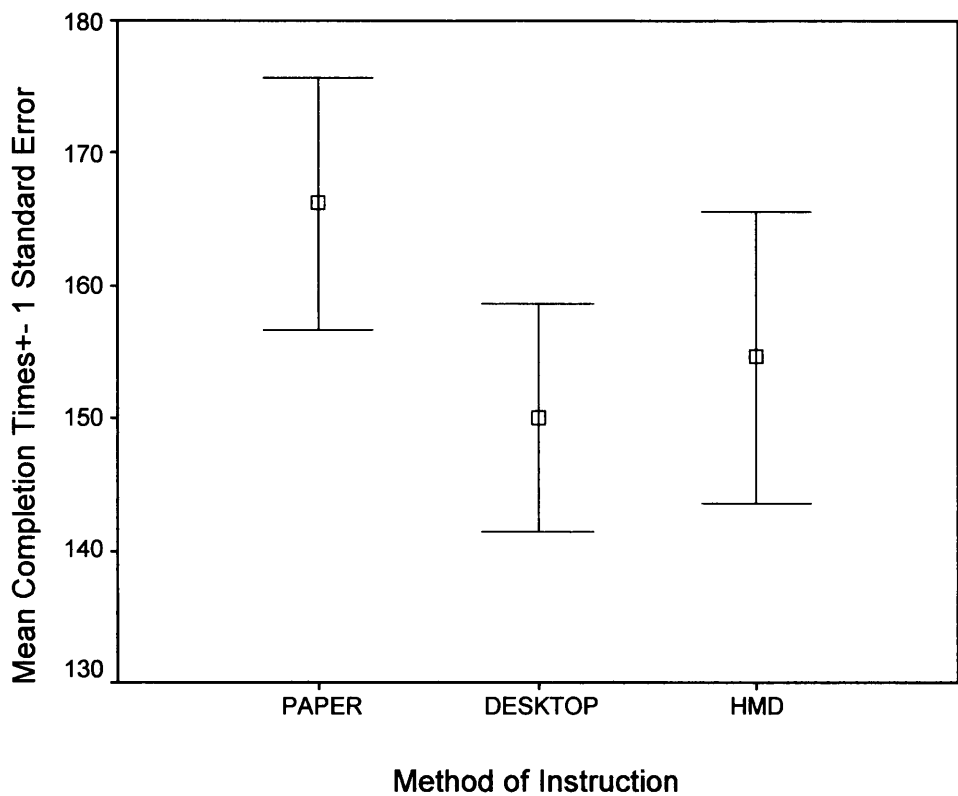


Figure 3.4 illustrates the mean completion times and plus or minus 1 standard error for the three models in each instruction condition. The paper instruction condition had the highest mean completion time of 166.2 seconds. The head-mounted display condition had the second highest mean completion time of 154.5. The desktop instruction condition had the lowest mean completion time of 150.2. The standard error for the paper and desktop are 9.5 and 8.6 seconds respectively. The standard error for the head-mounted display condition is 11.2 seconds. To test these differences further, the data were analysed using a one-way analysis of variance for repeated

measures. There was no significant effect for type of instruction $F(2,40)=0.962, N.S$ (See Appendix 3.8).

Table 3. 2: Mean error rates and standard deviations (in brackets) for each instruction condition: Paper, Desktop and Head-Mounted Display.

Paper Condition.	Desktop Condition.	HMD Condition.
0.9(1.2)	2.0(2.6)	2.1(3.0)

Table 3.2 shows the mean number of errors made in each of the three conditions. The paper condition had the lowest mean number of errors with 0.9. The desktop and head-mounted display had similar mean errors with 2.0 and 2.1 mean errors respectively. Standard deviations for the three conditions were 1.2 for the paper condition, 2.6 for the desktop condition and 3.0 for the HMD condition. To test these differences further, the data were analysed using a one-way analysis of variance for repeated measures. There was no significant effect for type of instruction $F(2,40)=2.431, N.S$ (see Appendix 3.8).

3.2.3 Qualitative Results.

This study also had a quantitative section that involved a semi-structured interview with each participant immediately after he or she had completed the test battery. The questionnaire had seven questions about the study, about the mechanics of the interactive programme and the instruction formats in the three experimental conditions. The results of these interviews are summarised in table 3.3.

Table 3.3: Qualitative feedback for Experiment 1.1

Question 1: Which figure did you find hardest to build?		
Figure	Frequency	Percentage
Beta	10	19
Gamma	4	19
Same	13	61.9
Question 2a: Which of the instruction formats, paper, desktop or head-mounted display did you find easiest to use.		
Instruction Type	Frequency	Percentage
Paper	10	47.6
Desktop	4	19
HMD	7	33.3
Question 2b: Which of the instruction formats, paper, desktop or head-mounted display did you find hardest to use.		
Instruction Type	Frequency	Percentage
Paper	5	23.8
Desktop	7	33.3
HMD	9	42.9
Question 3: In the desktop and head-mounted display conditions did you find the speed of the animation either too slow or just right?		
Speed	Frequency	Percentage
Too fast	10	47.6
Too slow	3	14.3
Just right	8	38.1
Question 5: Did you have any difficulty in identifying the size of the bricks.		
Difficulty in identifying	Frequency	Percentage
Yes	7	33.3
No	14	66.7
Question 6: Do you wear glasses or contact lenses.		
Wear eye correction	Frequency	Percentage
Yes	4	19
No	17	81

The first question asked which of the models the participants found the most difficult to build. This was to ascertain whether the models did have similar characteristics and that no model was harder to construct than the others. Of the twenty-one participants, four (nineteen percent) found the Gamma model the hardest, four found (nineteen

percent) the Beta the hardest, but thirteen (over sixty-one percent) thought that the three models had the same level of difficulty.

Question two was in two parts, in the first part the participants were asked which instruction format they found the easiest to use, ten (forty-seven and a half percent) said paper instructions, four (nineteen percent) said desktop and seven (over thirty-three percent) said head-mounted display. In the second part of question two the participants were asked which instruction format they found the hardest to use. Five (nearly twenty-four percent) of the participants found the paper instructions the hardest, seven (over thirty three percent) thought the desktop was the hardest and nine (over forty-two percent) thought the head-mounted display condition was the hardest.

Question three was another question that attempted to investigate participants' views on the mechanics of the programme, namely the speed of the animation. The speed of the animation was one frame every 4.5 seconds. Ten participants (over forty-seven percent) thought that this was too fast, three (fourteen percent) of the participants thought the speed was too slow and eight (thirty-eight percent) thought it was just about right.

Question four asked participants whether they found the control system in the interactive programme easy to use. All participants thought that the control system was easy to use and had reported no difficulty with it.

Question five asked the subjects whether they had problems identifying the size of the bricks in any of the conditions in the study. Seven (over thirty-three percent) said they

had difficulty in identifying the size of the bricks. Two of those participants remarked that they did not bother to read the textual information on the brick, e.g. 2x4 for two by four and picked up the wrong brick. On the other hand fourteen (almost sixty-seven percent) of the participants had no difficulty in identifying the size of the bricks. One of this particular group of participants felt that this textual information helped them to choose the correct brick.

In question six the participants were asked whether or not they wore eye correction. Four (nineteen percent) of the participants wore eye correction all the time and the remaining eighty-one percent never wore eye correction.

The last question was an open-ended question about the experiment and asked participants about using the head-mounted display for building the models. Here participants gravitated to one of two positions regarding the proposed advantage of having to just move the eyes to switch attention between the construction area and the head-mounted display to do the task. Six participants found this switching attention very difficult. Some found they lost their place following the instructions looking up and down; they experienced difficulty with coordination and could not take in the instructions fast enough. Some felt that switching their attention between the head-mounted display and the task slowed them down. One participant commented that it was easier for them to move their head rather than their eyes. Another person in this group felt that their attention on the head-mounted display was split in three, between the image of the brick, the text and the task. On the other hand eight participants liked the feature of moving the eyes between the screens in the head-mounted display and the task. Conversely they found an advantage in not having to move their head and

one felt that this system speeded up their performance in constructing the model. Two participants had no further comments to make about the head-mounted display. One participant thought the projected image on the head-mounted display looked far away. Another thought that more control of the programme was needed for the head-mounted display condition especially a pause function.

3.2.4 Discussion.

The results from the first experiment do not reveal any significant differences in mean completion time and errors rates presenting instructions on the opaque binocular head-mounted display, on a desktop computer or on paper. This is a different result to one previous study that compared monocular displays to more a conventional format. Ockerman, Najjar, & Thompson (1997a) compared teaching origami to participants on a wearable computer system with a learner centred locus of control to conventional book instructions. The book was significantly faster than the wearable computer system but more mistakes were made using the book than the wearable system, however the latter measure was not tested for significance. In another study (Barber et al., 1999) found that following instructions to solve a puzzle was slower using a monocular display than a desktop computer. However the latter researchers did not perform significance testing on the mean completion times in the conditions of their experiment. The puzzle task in the latter experiment is different from an assembly task. These differences in the result of experiment 1.1 to previous studies may be caused either by locus of control, type of head-mounted display or the type of task.

The only notable quantitative difference in experiment 1.1 was that participants took marginally longer to complete the paper instructions and made fewer errors in this condition. This may be indicative of a speed accuracy trade off; Wickens (1992) describes the completion time and error rate as two dimensions of the efficiency of processing information. An increase in the speed will cause a decrease in task accuracy and vice versa. Although attempts were made in this experiment to make the instructional programme as simple to operate as possible, the control the user has over the paper booklet is very different to that over the instruction programme. In the paper instruction condition participants could work through the instructions at their own pace whereas in the computer-based conditions they were working at the speed of the animation. As well as this the paper condition does not have the same issues with field of view as the computer-based instruction conditions. Another incompatibility is that users read a booklet of instructions in a different way than they view instructions on a screen. Users have a physical interaction with the paper instructions therefore they exert a different type of control whilst using paper instructions. These considerations resulted in the omission of a paper condition in the second experiment in this series of experiments.

Feedback from the interviews indicated that a significant number of the participants thought that the models were of similar characteristics and had no difficulty in building any particular model. Half the participants thought the animation was too fast but everybody found the control system easy to use. The lack of a pause function was seen as an important omission to the control system. Several participants had problems identifying the size of the bricks, which may have been due in part to the pictorial representations of the bricks in the animations being unclear. These findings

prompted the replacement of the animated instructions programme with a video based instruction clip on a video player with a pause function for the next experiment. As in the pilot study the paper condition was the most popular for following the instructions. The head-mounted display was the least popular way to view the instructions. Despite this, two thirds of the participants reported having no difficulty in switching attention between the head-mounted display screen and the task whilst following the instructions and assembling the model

3.3 Experiment 1.2: Efficiency Between two types of HMD and Desktop Instructions for Assembling an Abstract Model.

This experiment investigated the effectiveness of two different types of head-mounted displays and desktop for following instructions whilst constructing Lego models. The rationale behind this experiment was to compare the effectiveness of head-mounted displays that had the display screens in different positions. Thirty-one participants had to construct abstract shapes using instructions in three different conditions, with a “Albatech Personal Monitor” head-mounted display, a “Sony Glasstron” head-mounted display and a desktop computer. The Albatech and the Glasstron head-mounted displays have the display screens in different areas of the user’s field of view. The Glasstron is positioned on the user’s forehead whereas the Albatech sits on the bridge of the nose. The eye movements involved in using the two systems are different. The Glasstron requires the wearer to move their eyes up and down whilst switching attention from the instructions to the task. The Albatech on the other hand requires the user to look past the instruction screen to the task. The instructions in all conditions comprised of a video of a pair of hands building each model with added textual information about the bricks. The reason for this shift from the basic animation in the previous experiment was based on the premise that the video would give clearer representations of the bricks and convey more action information necessary for a procedural task. This video clip was played on a computer utility, Windows Media Player, with a similar control system to a video tape recorder. This is a type of interface that would be more familiar to participants. The controls on this utility include a pause button, a function incorporated following feedback from the previous study. It was hypothesized that there would be a difference between the three conditions in mean completion times and error rates for building the models.

3.3.1 Method.

Design

This study employed the same design as the previous experiment; it had a quantitative section and qualitative section. The independent and dependent variables were the same as the first experiment.

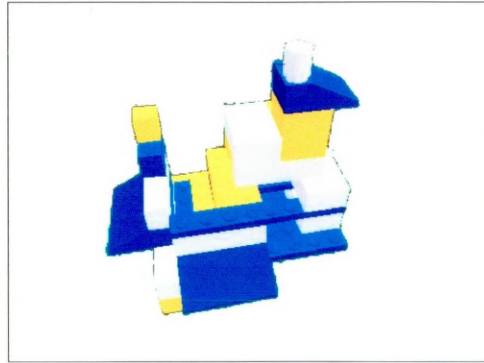
Participants

There were thirty-one participants, twelve males and nineteen females with an age range from eighteen to forty-six. The participants were students of the University of Abertay Dundee. All the participants were unpaid volunteers. All the participants had normal or corrected eyesight.

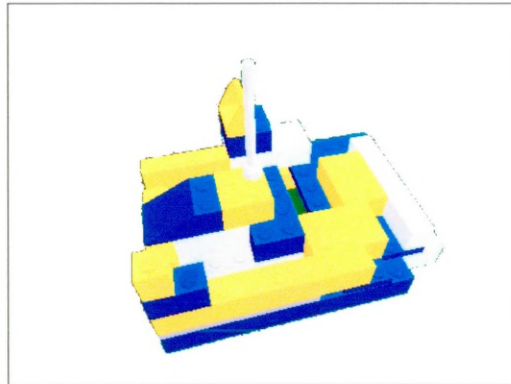
Materials and Apparatus.

The materials and apparatus used in this experiment were similar to the previous experiment except that an “Albotech” head-mounted display replaced the paper booklet in the test battery. As in the first experiment three abstract models of similar construction were used (See Figure 3.5). The models were called “Delta”, “Epsilon” and “Zeta” and had a different number of bricks to the previous experiment (twenty five in each model). The interactive computer programme to present instructions used in experiment one was replaced by a video clip of a pair of hands building the abstract model presented on “Windows Media Player”. Each model had a separate video clip

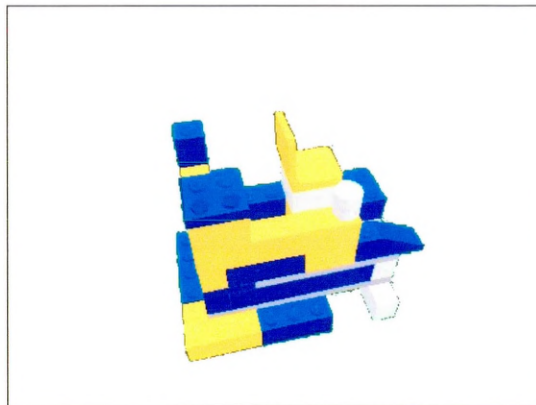
Figure 3.5: Abstract models used in Experiment 1.2



Delta

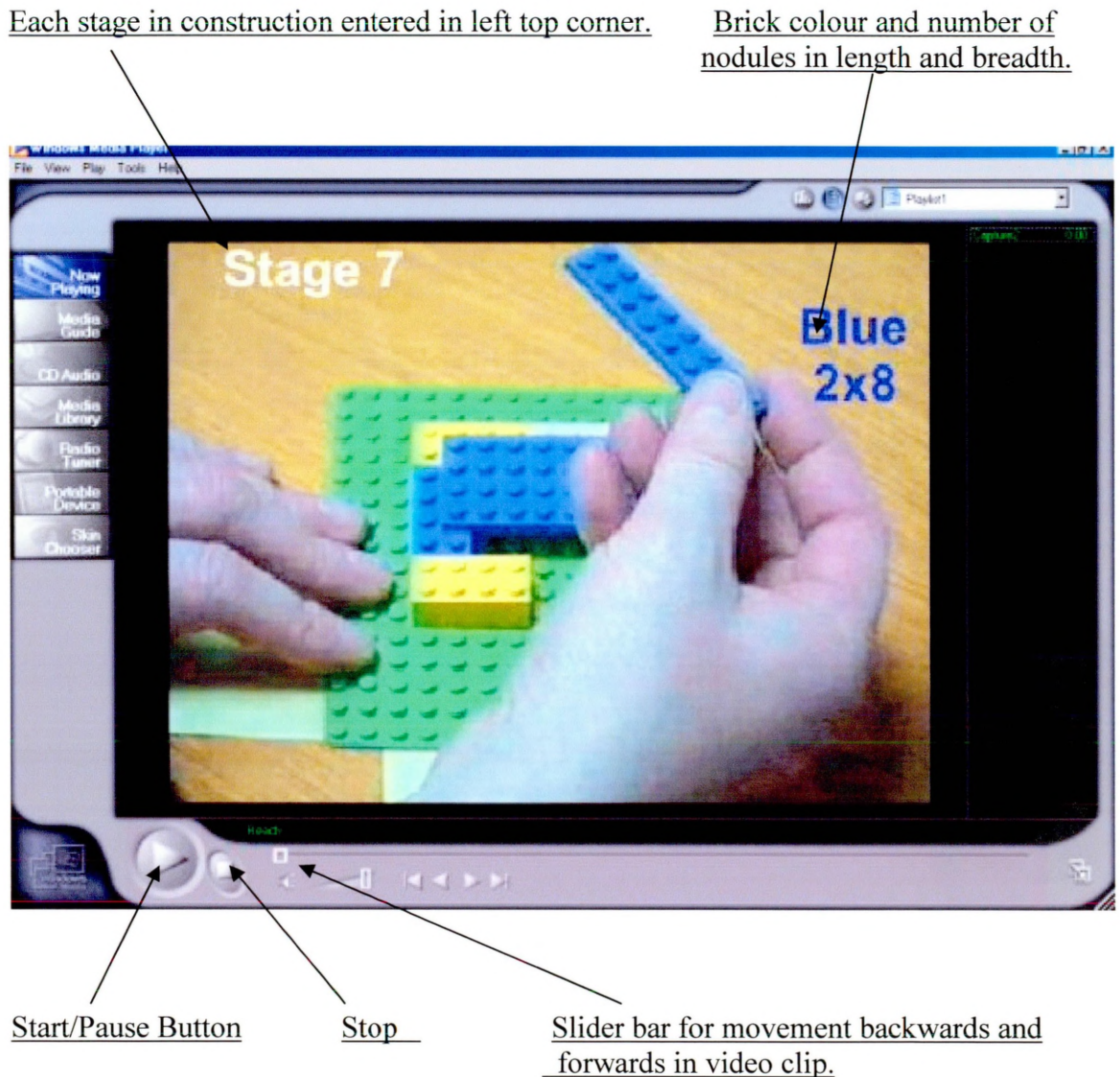


Epsilon



Zeta

Figure 3.6: Screenshot from Video clip used in Windows media player.



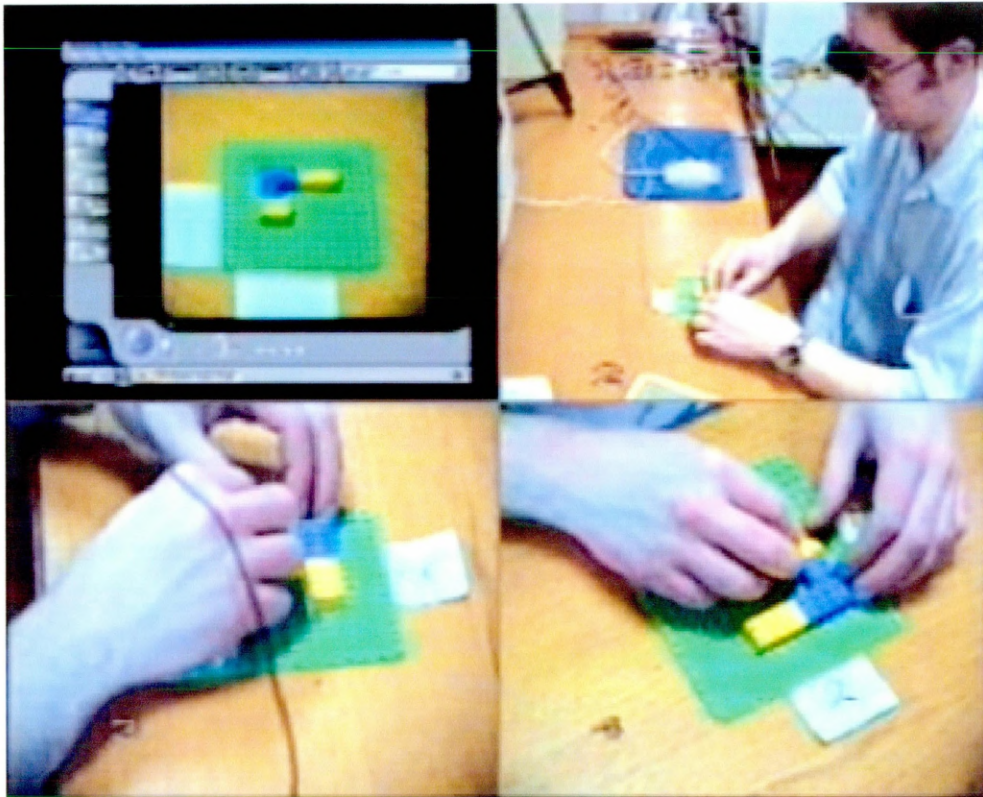
Participants followed a video clip of a pair of hands building the models. Each stage in the construction was placed in text in the left hand corner of the screen, this acted as a point of reference for the participants as they moved backwards or forwards in the video clip. The colour of the brick was placed in the right hand corner of the screen, in a text colour that was the same as the brick. The brick size was indicated by the number of nodules the brick was in length and breadth. For example 2X 8 represents the information “two nodules wide and eight nodules long”.

with text describing each brick in sequence required for the construction of the model (See Figure 3.6). In the qualitative part of the study participants were asked seven questions in a semi-structured feedback questionnaire similar to the previous experiment.

Procedure.

The Procedure for this experiment was very similar to experiment 1.1. The conditions in this experiment were slightly different; each participant followed the instructions via the desktop, Albatech or Glasstron. The counter-balancing procedure was the same and the participants were filmed building the abstract models in the testing bays. However in this experiment the experimenter explained the control system of Windows Media Player to the participants. The participants followed the pair of hands on the video clip building the model rather than the animation in the previous study. They now had the opportunity to pause the video clip and move backwards and forwards in a linear fashion. For the purpose of a more detailed analysis the participants were filmed by three cameras and the participants' use of the Windows Media Player programme was also filmed. This output was brought together onto one screen using CORIOscan, a utility that produces a four-way screen (See Figure 3.7). The performance of each participant was videoed from the four-way screen and then the performance of each participant was later analysed. The length of time it took each participant to build each model and the number of errors made was calculated. As in experiment 1.1 the participants were given a post-test semi structured interview after they had completed the experiment.

Figure 3. 7: Four way video analysis with CORIOscan



Screenshot of 4-way analysis of participants' performance of assembly task.

The four-way screen used for the analysis was created using CORIOscan, a specialist scan converter. This converter allowed images from three camera angles and the monitor screen to be placed on one television and video taped for further analysis.

3.3.2 Quantitative Results.

Figure 3.8: Graph Showing Mean Completion Times and Standard Error in seconds for each instruction condition, Albatech, Glasstron and Desktop.

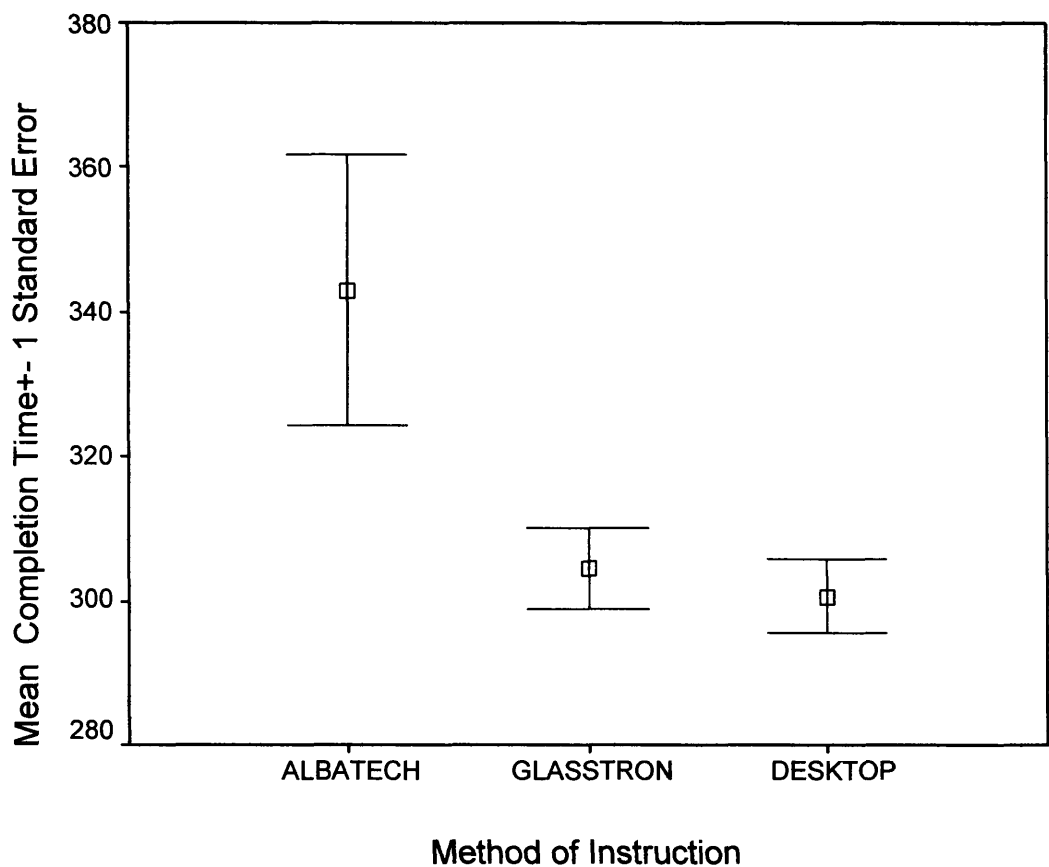


Figure 3.8 illustrates the mean completion times and plus or minus 1 standard error for the three models in each instruction condition. The Albatech instruction condition had the highest mean completion time of 343.1 seconds. The other two conditions had similar completion times to one another. The head-mounted Glasstron condition had the second highest mean completion time of 304.6. The desktop instruction condition had the lowest mean completion time of 300.7. The standard error for the Albatech and Glasstron are 18.7 and 5.6 seconds respectively. The standard error for the desktop condition is 5.2 seconds. To test these differences further, the data were

analysed using a one-way analysis of variance for repeated measures. There was a significant effect for type of instruction $F(2,60)=5.373$; $p<0.05$ (See Appendix 3.9). Pair-wise comparisons using paired sample t-tests were carried out at the 5 percent level on the mean completion times for the three conditions. To achieve significance with three conditions the t-tests must show significance beyond the 0.02 percent level. The only comparison of means that achieved significance beyond the 0.02 percent level was between the desktop and Albatech, $t(30)=2.851$, $p<0.02$.

Table 3.4: Mean error rates and standard deviations (in brackets) for each instruction condition: Monitor, Glasstron and Desktop.

Albatech	Glasstron	Desktop
2.71(3.09)	1.10(2.43)	0.84(2.43)

Table 3.4 shows the mean number of errors made in each of the three conditions. The monitor condition had the highest mean number of errors with 2.71. The desktop and the Glasstron had 0.84 and 1.10 mean errors respectively. To test these differences further, the data were analyzed using a one-way analysis of variance for repeated measures. There was a significant effect for type of instruction $F(2,60)=7.093$; $p<0.05$ (See Appendix 3.9). Pair-wise comparisons using paired sample t-tests were carried out at the 5 percent level on the mean completion times for the three conditions which required the t-tests to show significance beyond the 0.02 percent level. Two of the pair-wise comparisons achieved significance beyond the 0.02 percent level, first between the desktop and Albatech, $t(30)=3.627$, $p<0.02$., second between the Glasstron and Albatech, $t(30)=2.882$, $p<0.02$.

3.3.3 Qualitative Results

As with experiment 1.1 this experiment also had a qualitative section where a semi-structured interview was held with each participant immediately after he or she had completed the test battery. Table 3.5 indicates the results from some of the questions on the feedback questionnaire. The first question had two parts; participants were asked if they found that one of the models easiest to build. Although twenty-eight (ninety percent) felt that the models were much the same, three (just over nine percent) thought zeta was easiest. The second part of question one asked the participants which model they found the hardest to build. Two (six and a half percent) thought delta was the hardest, three (over nine percent) felt that epsilon was the hardest to build and twenty-six (nearly eighty-four percent) found no particular model harder than another to build.

Question two also had two parts. First participants were asked which instruction condition they found the easiest. Twenty-four (seventy-seven percent) thought the desktop was the easiest, one (just three point two percent) found the Albatech easiest and five (sixteen percent) the Glasstron. In part two of question two participants were asked which instruction condition they found the hardest. Twenty-seven (eighty-seven percent) found the Albatech condition the hardest. Two (six and a half percent) thought the Glasstron was the hardest and two thought they were both equally hard. None of participants thought the desktop was the hardest.

Table 3.5: Qualitative feedback for Experiment 1.2

Question 1a: Which of the models did you find the easiest to build?		
Model	Frequency	Percentage
Zeta	3	9.7
No difference	28	90.3
Question 1b: Which of the models did you find the hardest to build?		
Model	Frequency	Percentage
Delta	2	6.5
Epsilon	3	9.7
No Difference	26	83.9
Question 2a: Which instruction condition did you find the easiest to use?		
Instruction condition	Frequency	Percentage
Albatech	1	3.2
Desktop	24	77.4
Glasstron	5	16.1
No Difference	1	3.2
Question 2b: Which instruction condition did you find the hardest to use?		
Albatech	27	87.1
Glasstron	2	6.5
No Difference	2	6.4
Question 3: Was the speed of the instruction video too fast, too slow or just right?		
Speed	Frequency	Percentage
Too fast	3	9.7
Too slow	7	22.6
Just right	21	67.7
Question 5: Did you identify the size of the brick by its visual representation, text information about the size of the brick or by both.		
Mode of identification	Frequency	Percentage
Visual	6	19.4
Text	10	32.3
Both	15	67.7
Question 6: Did you have problems with switching your attention between the instruction screen and the task in the head-mounted display conditions.		
Difficulty with switching	Frequency	Percentage
Yes	10	32.3
No	21	67.7
Question 7: Which head-mounted display did you find switching attention the easiest?		
Head-mounted display	Frequency	Percentage
Albatech	8	25.8
Glasstron	22	71
No difference	1	3.2

The third question asked the participants how they felt about the speed of the video and. A significant number of participants, twenty-one in all (nearly sixty-eight percent) felt that the speed was okay. Only three (almost ten percent) thought that the animation was too fast and seven (over twenty-two percent) thought it was too slow. The fourth question asked if the control system was easy to use. All participants in the study reported that the control system was easy to use.

In the fifth question participants were asked if they identified the size of brick by its visual representation, textual information about the size of the brick or both. Six (nineteen and a half percent) identified the brick by visual information alone, ten (just over thirty-two percent) identified the brick by text alone and fifteen (nearly forty-eight and a half percent) identified the brick by both text and visual representations.

A sixth question asked if the participants had problems with switching their attention between the instruction screen and the task in the head-mounted display conditions. The majority, twenty-one (over sixty-seven percent) reported no difficulty in doing so; however the remaining ten, nearly a third had problems switching attention.

The seventh and last question asked the participants which head-mounted display they found easiest to switch attention in. Eight found the Albatech the easiest (over twenty-five percent), only one (three point two percent) found both the same however a significant number, twenty-two (seventy-one percent) reported that they found the Glasstron easiest to use in this manner.

As part of the last question participants were also asked if they had any further comments about the experiment. Most of them reported problems viewing the video clip using the Albatech. Some complained that the screen appeared blurred; others found the information on the video difficult to discern. A number found the method of adjusting the mirrors to get a single image very difficult. There were also issues with this headset fitting on user's heads. Some participants complained that the front was too heavy and kept slipping forward.

3.3.4 Discussion

The results from this experiment reveal that the Albatech head-mounted display was significantly slower than the desktop and less accurate for following the instructions than the Glasstron or the Desktop. The feedback from the semi-structured interviews show that the participants found the Albatech hard and problematic to use. The display screen component clips on a pair of eyeglasses that are provided as part of the head-mounted display, alternatively the display screen component can be clipped onto the glasses normally worn by the user. The weight lies to the front of this screen display component and on occasion the eyeglasses were pulled forward by the weight. One ergonomic issue encountered with the Albatech is that the eyeglasses did not properly fit every participant's head. Another usability issue was that each participant had to adjust the two mirrors on the Albatech until they could see one image. Some users may not have adjusted the mirrors properly in this experiment, which would result in viewing a blurred or double image.

The display itself can only use a television or video signal whereas the Sony Glasstron uses a SVGA signal which projects a much sharper image. Another

disadvantage for the Albatech it has a smaller field of view than the Glasstron. The details of the bricks and text instruction are degraded and the image is blurred in the Albatech screen. These factors would contribute greatly to the poor performance of the Albatech in this experiment relative to the other conditions. The dissimilarities between the physical usability of the Albatech and the Glasstron and desktop meant that a comparison in performance between the Albatech and the other two conditions was problematic.

As in the first experiment approximately two thirds of the participants had no difficulty switching attention in two head-mounted display conditions. However about three times as many participants found switching attention easier using the Sony Glasstron than the Albatech. This would suggest that a binocular head-mounted display with the display occupying the top half of the visual screen might be favoured more by users than an image projected in the centre of the visual field. Findings concerning the representations participants used to identify the bricks also mirrored results from the first experiment. A majority used both the text and visual information to identify each brick in the assembly.

3.4 Experiment 1.3: Comparison between two Modes of using a Head-mounted Display and Desktop Instructions for Assembling an Abstract Model.

This study compared two modes of using a head-mounted display and desktop instructions in constructing models using children's play bricks. Twenty-seven participants constructed abstract shapes using instructions in three different conditions; with a Sony Glasstron in "transparent mode", a Sony Glasstron in "opaque mode" and a desktop computer. In the two Glasstron conditions the participants used different eye movements to follow the instructions and complete the tasks. In the see through mode the instruction screen was transparent so users could look through the screen and see their hands and the play bricks beyond. Using the Glasstron in this mode the participants shift the focus of their eyes from the instructions to the task. In the opaque mode participants switch their gaze back and forward from following the instructions on the screen to building the models.

As in the previous experiment the instructions in all the conditions comprised of a video showing a pair of hands constructing each model. However in experiment 1.3 the video clip had a vocal soundtrack giving simultaneous information about the size, colour and shape of each brick rather than on screen text information. The vocal track was used because text may be difficult to read in the transparent or see-through mode on the Glasstron. The voice was played through the computer speakers rather than the headset due to the difficulty in adjusting and controlling the sound in the earpieces on the Glasstron. Each participant would have to control the sound level himself or herself, this may have produced a wide range of sound levels heard by participants.

3.4.1 Method.

Design

This study employed the same design as the previous experiments it had a quantitative section and qualitative section. The independent and dependent variables were the same as the previous two experiments.

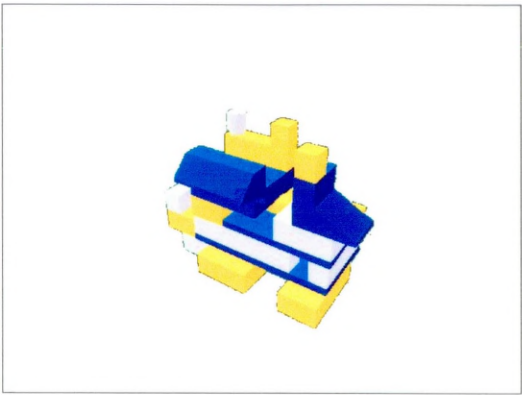
Participants.

There were twenty-seven participants; thirteen males and fourteen females with an age range from eighteen to forty-five. The participants were students of the University of Abertay Dundee. All the participants were unpaid volunteers. All the participants had normal or corrected eyesight.

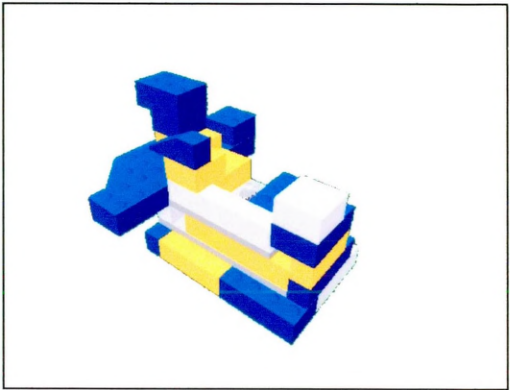
Materials and Apparatus.

The materials and apparatus used in this experiment were almost the same as the previous experiment except that the “Albatech” head-mounted display was replaced by the Glasstron used in “transparent” mode. As in previous experiments three abstract models were used, each with the same number of bricks as the previous experiment (twenty five in each model) and a similar construction but this time

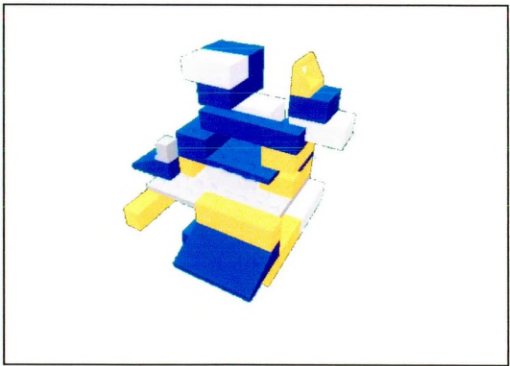
Figure 3.9:Abstract Models used in Experiment 1.3



Eta



Iota



Theta

they are called “Eta”, “Iota” and “Theta” (See Figure 3.9). As in the previous experiments the instructions comprised of a video clip of a pair of hands building the abstract model presented on “Windows media player”. However the video clip had a vocal narration that gave each stage of the construction a description of the brick needed for the stage. This vocal narration replaced the text descriptions used in the previous experiment. In this study the vocal instructions emanated from speakers attached to the same computer as the Glasstron was connected to. As in the previous experiments participants were asked seven questions in a semi-structured feedback questionnaire. The questions were the same as the previous experiment.

Procedure.

The procedure of this experiment was almost identical to the previous experiment, with the same type of counter-balancing for allocating the participants in the three conditions was used. The use of the testing bays and recording equipment was the same as the previous experiment. However in the see-through Glasstron condition 60 watts of additional light was shed on to the area where the models were being constructed in order that participants could clearly distinguish the real environment through the head-mounted display.

3.4.2 Quantitative Results.

Figure 3.10: Graph showing mean completion times and standard error in seconds for each instruction condition: Glasstron Transparent, Glasstron Opaque and Desktop



Figure 3.10 illustrates the mean completion times and plus or minus 1 standard error for the three models in each instruction condition. The Glasstron transparent condition had the highest mean completion time of 231.9 seconds. The desktop condition had the second highest mean completion time of 227.2. The Glasstron opaque condition had the lowest mean completion time of 225.1. The standard error for the Glasstron transparent and Glasstron opaque are 6.6 and 3.5 seconds respectively. The standard error for the desktop condition was 3.6 seconds. To test these differences further, the data was analysed using a one-way analysis of variance for repeated measures. There

was no significant effect for type of instruction $F(2,52)=0.641$; NS.(See Appendix 3.10).

Table 3. 6: Mean error rates and standard deviations (in brackets) for each instruction condition: Monitor, Glasstron and desktop.

Glasstron Opaque.	Glasstron Transparent	Desktop Condition.
1.2(1.5)	2(2.2)	1.6(1.7)

Figure 3.6 shows the mean number of errors made in each of the three conditions. The Glasstron transparent condition had the highest mean number of errors with 2. The desktop had the second highest mean errors with 1.6 and the Glasstron in opaque mode had the lowest mean number of errors with 1.2. To test these differences further, the data were analyzed using a one-way analysis of variance for repeated measures. There was a no significant effect for type of instruction $F(2,52)=2.246$; NS. (see Appendix 3.10).

3.4.3 Qualitative Results

As with the previous experiments there was an interview each participant immediately after he or she had completed the test battery. This interview had seven questions about the study, about the mechanics of the instruction video and the instruction formats in the three experimental conditions. A number of the more salient questions are shown in table 3.7

Table 3. 7: Qualitative feedback for Experiment 1.3

Question 1a: Which of the models did you find the easiest to build?		
Model	Frequency	Percentage
Iota	1	3.7
No difference	26	96.3
Question 1b: Which of the models did you find the hardest to build?		
Model	Frequency	Percentage
Eta	3	11.1
Iota	3	7.4
No difference	22	81.5
Question 2a: Which instruction format did you find the easiest to use?		
Format	Frequency	Percentage
Transparent Glasstron	1	3.7
Opaque Glasstron	9	33.3
Computer	1	63.0
Question 2b: Which instruction format did you find the hardest to use?		
Format	Frequency	Percentage
Transparent Glasstron	19	70.4
Opaque Glasstron	4	14.8
Computer	3	11.1
No difference	1	3.7
Question 3: What the speed of the video presentation too slow, too fast or just right?		
Speed	Frequency	Percentage
Too fast	3	11.1
Just right	24	88.9
Question 4: Did you find the control easy or hard to use?		
Control	Frequency	Percentage
Easy	5	18.5
Hard	1	3.7
Okay	1	3.7
Not used	20	74.1
Question 5: Did you identify the size of the brick by its visual representation, vocal description about the size of the brick or by both.		
Mode of identification	Frequency	Percentage
Visual representation	2	7.4
Vocal description	11	40.7
Both	14	51.9
Question 7: Which head-mounted display did you find easier to switch attention?		
HMD	Frequency	Percentage
Glasstron transparent	5	18.5
Glasstron opaque	20	74.1
No Difference	2	7.4

The first question was in two parts. In the first part, participants were asked if they found any of the models easier to build than the others. A majority of the participants,

twenty-two (over eighty-one percent) thought the models were the same. Three thought Eta was the easiest and two thought Iota was the easiest. The second part of question one asked participants if they found any the models harder to build than the others. As in the previous question a significant number, twenty-six (over ninety-six percent) there was no difference in difficulty between the models. One participant thought Iota harder than the other two.

Question two also had two parts; the first part asked the participants which instruction format they found easiest. The result showed that a majority of participants, seventeen altogether (sixty-three percent), found using the desktop the easiest. One liked the see-through Glasstron and the other nine (thirty-three percent) preferred the Glasstron in opaque mode. The second part of question two asked the participants which instruction format they found the hardest to use. Most, nineteen (seventy percent) thought that the Glasstron in transparent mode was the hardest. Four thought the Glasstron in opaque mode the hardest, three thought the desktop was the hardest and one felt that they were all equally hard to use.

In question three participants were then asked how they felt about the speed of the video, twenty-four (nearly ninety percent)thought the speed was okay. Three thought the speed of the video was clip too fast.

The participants were asked in question four if they used the control system and if so did they find it easy to use. A large number (seventy-four percent) did not use the control system, the few that did use the system, one participant most found it hard to use, the remaining five (eighteen and a half percent) found it easy to use.

The fifth question asked participants if they identified the size of brick by its visual representation, spoken information about the size of the brick or both. Fourteen or over fifty percent of participants identified the size of brick both by vocal and visual representation together. Eleven participants (over forty percent) used the spoken representations and two used visual representations alone.

In question six the participants were asked if they had problems with switching their attention between the instruction screen and the task in the head-mounted display conditions. Twenty-one (nearly seventy-eight percent) reported no difficulty in doing so; however six participants, nearly a quarter, had problems switching attention in the two head-mounted display conditions.

Question, number seven, asked the participants which head-mounted display they found easiest to switch their focus from the instruction screen to the task. Twenty (seventy-four percent) reported that they found the Glasstron in opaque mode easiest to use in this manner. Five found the Glasstron in see-through mode easier and two participants felt they were both the same for switching gaze. They were also asked if they had further comments about the head-mounted display conditions in the experiment. Some thought that the screen in the see-through Glasstron condition was too busy with the instructions superimposed on the task. They further reported difficulties in extracting details from the instructions and building the models at the same time in the see through condition. Some participants found the see-through mode hard to use if they moved their heads since the instruction screen moved around

the room in their visual field. This feature of the see through mode created competing moving images in their visual fields.

3.4.4 Discussion

In the third experiment there were no significant differences between the conditions in mean completion times or error rates for the models. The Sony Glasstron in transparent mode had a higher mean completion time and higher error rates than the other two conditions but these differences were not significant. Feedback from the semi-structured interviews showed that most participants used both the spoken and visual information about the bricks as a guide to help them to build the models. Most participants found it easy to switch their gaze from the instruction screen to the task. When asked which head-mounted display they found easiest to use and switch attention in, most reported that the Sony Glasstron in the opaque mode was the easiest to use and for switching attention from the display to the task. Many participants found that performing the experimental task with the instructions on the see-through Glasstron screen produced a very busy visual field. Picking out details from the instructions and building the models at the same time in the see-through mode was difficult. Some participants found that when they moved their heads, the instruction screen moved around the room in their visual field. This caused them to lose sight of their hands building the model. This negative feedback for the see-through screen is consistent with studies using HUDs, for example (Sojourner & Antin, 1990).

3.5 General Conclusions for first phase of experiments.

Table 3.8: Summary of Quantitative results of experiments in Phase 1

Experiment	Conditions	Findings
Experiment 1.1	<p>Instructions on Paper</p> <p>Instructions on Desktop P.C.</p> <p>Instructions on Opaque Glasstron</p>	No significant difference between the conditions in both dependent variables; time and errors.
Experiment 1.2	<p>Instructions on a Desktop P.C.</p> <p>Instructions on Opaque Glasstron</p> <p>Instructions on Albatech Monitor.</p>	<p>No significant difference between the Desktop and Opaque Glasstron conditions in both dependent variables.</p> <p>Albatech Monitor condition significantly slower than other conditions and significantly more errors made in Albatech condition.</p>
Experiment 1.3	<p>Instructions on Desktop</p> <p>Instructions on Opaque Glasstron</p> <p>Instructions on See-through Glasstron</p>	No significant difference between the conditions in both dependent variables; time and errors.

Table 3.8 illustrates the main findings of the quantitative research in the experiments undertaken in phase one. The implications of the results of experiment 1.1 for the first research question are that when animated instructions are used for following a task based on programme control, there is no significant difference in efficiency between a Glasstron head-mounted display and traditional methods of instruction. This finding is

different to the few studies comparing the performance of monocular displays to other instruction methods.

Results from experiment 1.2 revealed that the Albatech had a significantly slower completion times in the task in comparison to the desktop and significantly poorer accuracy for the task compared to the desktop and the Glasstron. However this could be attributed to reasons other than efficiency in building the models and following instructions on a screen in the middle of the user's field of view. One reason may have been the poorer image this head-mounted display projected compared to the SVGA signal used in the other conditions in the experiment. Another shortcoming of the Albatech was that the participants themselves had to adjust the inter screen distance on this head-mounted display which may not have been done correctly. An incorrect adjustment may lead to viewing a distorted or blurred image. The implications of these findings were that the second research question could not be answered due to these above technological differences. Comparisons between these particular head-mounted displays could not be made.

There was no significant difference in task performance between the Glasstron in opaque mode and the Glasstron in the see-through mode. Participants were able to follow the instructions just efficiently using the Sony Glasstron to look through the instruction screen at the task as they were switching focus between the instruction screen and task using the Glasstron with the opaque screen. As far as the third research question is concerned these findings indicate no quantitative differences between see-through and opaque displays for following assembly instructions.

The general conclusions from the qualitative sections of the studies in the first phase of experiments can be summarised by the following paragraphs. The results illustrated that most participants used both visual and verbal information to build the abstract models. Although most participants reported no difficulty in switching attention between the instruction screen and task, up to a third in one experiment said they had difficulty following the instructions on the head-mounted display whilst building the models. Most participants found it easier to switch attention from down from the screen to the task in the opaque Glasstron than past the screen in the Albatech Monitor condition. Feedback from the participants indicated that they had problems using the Glasstron in transparent or see through mode. Looking through the transparent instructions to the task was difficult but task performance in this condition, although poorer was not significantly different from using the Glasstron in its opaque mode.

One original intention of the experiments was to develop a comparison between different subtypes or designs of head-mounted displays. Specifically these experiments were designed to ascertain if eye-movements between instructions and task were more efficient in a particular design. Technological differences between the head-mounted displays made available for the research programme rendered such comparisons impossible. A second phase of experiments concentrated of the fourth research question, finding the optimum configuration of text, sound and visual instructions on a video demonstration of an assembly task.

Chapter 4

4 Focus of the second phase of research programme.

“The emergence of computer-driven “hybrid” technologies has spawned unprecedented interest, yet advances in technological capability alone no more improve instruction than sharpened pencils prose.”

Hannafin & Rieber (1989).

4.1 Factors that effect comparisons between head-mounted displays.

This chapter explains the reasons for the focus of the experiments in the second phase of experiments in the research programme. Primarily this chapter will explore the various factors that militate against a comparison between different subtypes of head-mounted displays. A majority of these factors concern visual and physical individual differences in the human population. The latter dissimilarities would necessitate a greater degree of adjustability in the various headsets that were available for testing before any comparison could be made. Most of the sections in this chapter will list the usability and visual problems encountered in phase one when comparisons were made between the Albatech and the Sony Glasstron head-mounted displays. There will also be a section describing the evolving nature of head-mounted display technology.

One original objective of the research programme was to discover if there was an optimum method of displaying multimedia training materials on different types of head-mounted displays and which display type or design was the most advantageous

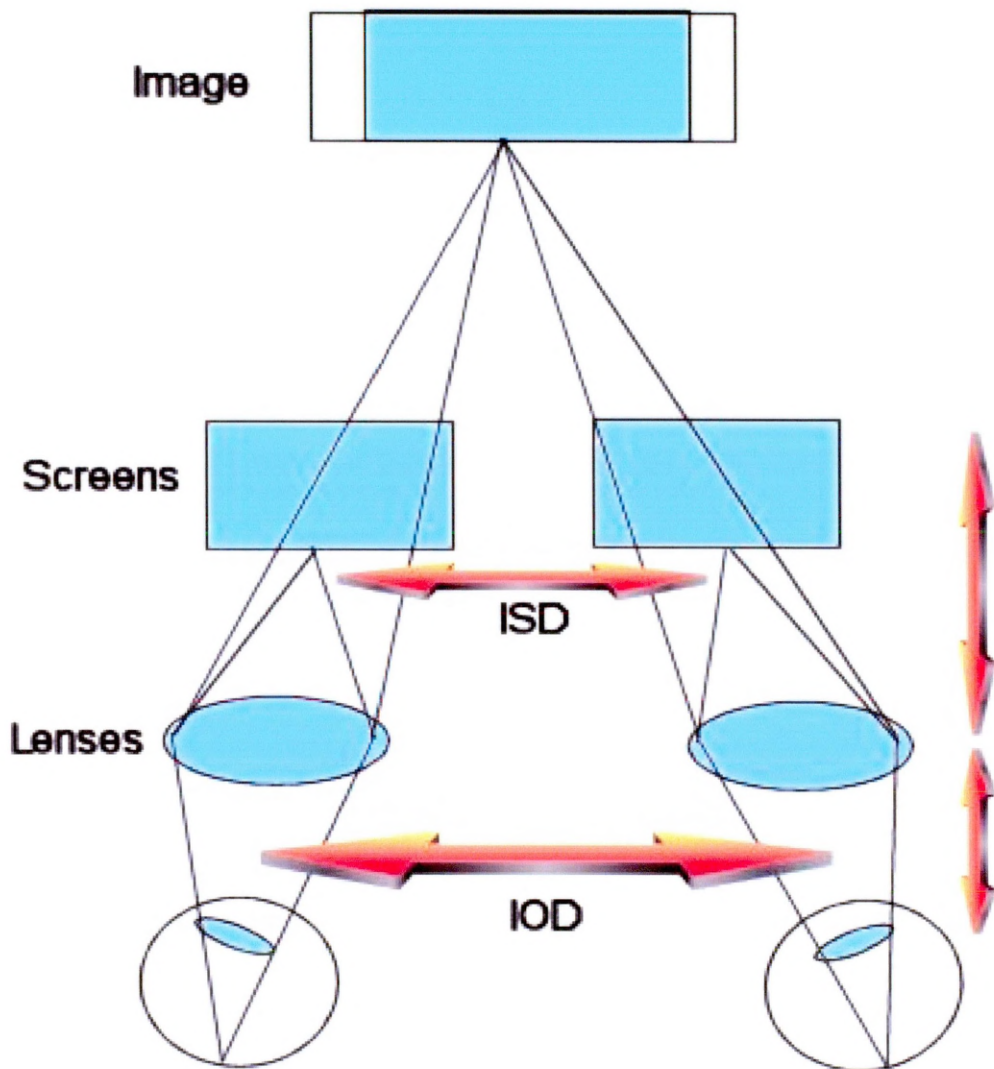
for following animated multimedia instructions for a procedural task. Results from the first phase of experiments uncovered several unforeseen problems with comparing the effectiveness of head-mounted displays for following multimedia instructions to traditional methods and comparing the effectiveness of different types of head-mounted displays. There were issues with the eyesight of the participants, not just about whether they wore their eye correction but viewing screens so close to the eyes could affect task performance since there are large individual differences in eyesight. As well as visual problems there are difficulties with the adjustability of the head-mounted displays to allow for the individual differences in the physical properties of user's heads. The headsets used in this research had different levels of adjustability. There exist a number of visual and physical differences that effect comparisons of head-mounted displays commercially available on the market.

4.1.1 The existence of individual differences in Interpupillary Distances.

One aspect of binocular head-mounted displays that can affect the vision of most users is the distance between the eyes relative to the distance between the lenses or screens in the head-mounted display, more specifically the difference between the users' interpupillary distance (IPD) and the optical centre of the head-mounted display (Piantanida, 1993). The design of most binocular displays consists of a pair of screens with an arrangement of mirrors or lenses positioned in front (See figure 4.1). In the general population there exists a variation in interpupillary distance.

According to Howarth (1999) IPDs range from 53-73 mm with an average IPD for the adult population of roughly 63mm. Few head-mounted displays on the market

Figure 4.1: Principal factors in Head-Mounted Display Design.

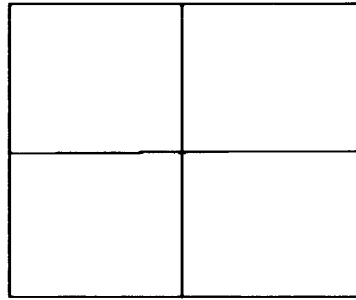


Arrows show possible parts of a binocular head-mounted display that could be adjusted to suit each individual user. The inter-screen distance (ISD) and the inter-ocular distance (IOD). (Adapted from Howarth 1999).

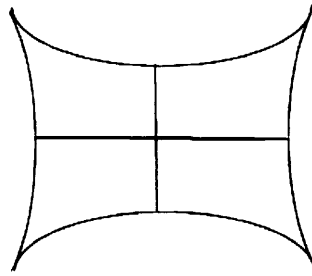
have adjustable lenses or screens that ensure that the optical centre can be brought into line with the user's eyesight. However a problem with adjustable systems is that they may cause optical distortions if not adjusted properly by the user. If the user wears a head-mounted display with the inter-screen distance different to the distance between the user's eyes then image deviation occurs, this may result in an increased demand on the user to fuse the images (Melzer & Moffit, 1997). The Glasstron could be adjusted to fit most heads but had no means of adjusting the distance between the two screens the users looked into. The Albatech had no adjustability for the size of a user's head but could be adjusted to different inter-screen distances to cater for individual differences in the distance between the user's eyes or interpupillary distances.

The lens systems of many HMDs produce distortions, for example squares seen through the lens may be distorted into pin cushion or barrel shapes (see figure 4.2). Such distortions become greater the more distant the eye is from the optical centre of the lens system. As far as the user is concerned items seen through the centre of the lens do not deviate from the centre but items viewed off centre appear dislocated as they were seen through a prism. If the fixed lenses in the head-mounted display have optical centres that are separated by more than the user's IPD (interpupillary distance) then the user will experience an outward divergence of the right and left eye images. Conversely, if the fixed lenses in the head-mounted display have optical centres that are separated by less than the user's IPD then the user will experience an inward divergence of the eyes when looking through the lens system. To look through the displaced optics of the head-mounted display both eyes may have to converge

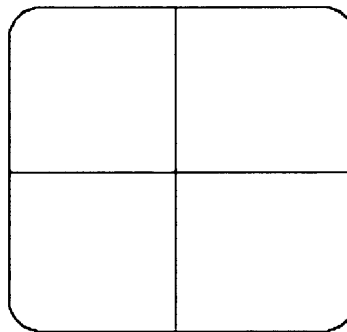
Figure 4. 2: Visual Distortion in Head-mounted Displays.



O



A



B

O represents viewed image, A pincushion distortion and B barrel shaped distortion.
(Adapted from(Millodot, 2000)).

or rotate inwards. For viewing something distant the user would have to rotate their eyes more in this head-mounted display than if viewing the same scene in the real world. Initially this over-convergence may be uncomfortable but the human visual system is very adaptive and it can quickly readjust the extraocular muscles of the eyes so that the vergence movements allow for the optical distortions of the unadjusted head-mounted display. However, when the viewer's eyesight has adapted to a head-mounted display with unadjusted optics the viewer would become slightly esophoric or cross-eyed.

The eyesight of humans is either orthophoric, that is the eyes point directly at fused binocular targets or slightly exophoric; that is the eyes point slightly outward. Most people with normal vision are slightly exophoric. Using a head-mounted display with fixed optics causes many users to demonstrate an exophoria that happens quickly then stabilizes as the vergence system adjusts to compensate for the prismatic deviation. When taking off the head-mounted display users would experience visual discomfort a second time because their vergence system would be adapted to the head-mounted display optics rather than to the real environment (Piantanida, 1993).

Of the two head-mounted displays used in the first phase of experiments, only one, the Albatech could be adjusted for interpupillary distance. This head-mounted display is adjusted by moving the mirrors in front of each eye. The range of this adjustment was between 58 mm and 77mm. A major drawback of this type of adjustable system is that the user may not line up the mirror or lenses exactly in front of both eyes thus causing the user to view a blurred or double image. The Glasstron does not have the facility of altering the optics to allow for differing interpupillary distance. This

difference in optical display settings between head-mounted displays constitutes an important obstacle for comparing different head-mounted displays. Users in general are not be experienced enough to make the various adjustments necessary and would require training.

4.1.2 Problems associated with field of view and screen image.

Other problems arose due to the differences in the projected images provided by the displays investigated. The two head-mounted displays, the Albatech and the Glasstron project screen images of different clarity, unequal distances and different fields of view to the user. These disparities in hardware performance meant that it was difficult to tell if the results were due to the characteristics of the viewed image or the display screen position on each head-mounted display. The results from the first phase could have been caused by the fact that the Albatech used a video signal and displayed a poorer image than the Glasstron , the latter used a much clearer SVGA output from a computer. These headsets could not be adjusted to equalise the visual differences between them and it was not possible to acquire different types of head-mounted displays that had the same screen configurations to continue comparisons.

The head-mounted displays that were made available for the research in this thesis represented a cross section of design types available on the market. Altogether there were five headsets that were available to be used for experimentation. Four of these head-mounted displays were binocular and the fifth was monocular. The designs of two of the binocular head-mounted displays were such that neither could be used in the experiments. One of these head-mounted displays was fully enclosed around the eyes thus removing all periphery vision when worn. This precluded engaging in a

secondary task whilst using this headset. The second head-mounted display had a design that prohibited users wearing eyeglasses for eye correction. This latter headset had to be omitted from the first phase of research since using them would exclude a sizable proportion of the potential participant population. This left two binocular head-mounted displays, the Glasstron and the Albatech and a monocular display, available for use in the studies.

These two binocular headsets were used in the first phase of the experiments that investigated the effectiveness of head-mounted displays for following multimedia instructions on a head-mounted display. The second experiment in this phase of experiments compared the efficiency of the Glasstron to the Albatech and a desktop computer for following the multimedia instructions for a procedural task. The results from this study showed the Albatech to be poorer in speed compared to the desktop P.C. and have a higher error rate compared to the Glasstron and desktop. It was concluded from post-test interviews that the video signal the Albatech used may have affected the participants' performance in the experimental task since participants found it harder to discern the features of the display compared to the SVGA image displayed by the Glasstron. Some participants reported that the displayed image on the personal monitor looked more distant and blurred. This experiment used text as a medium to convey information on the instructional video clip that was difficult for participants to read on the Albatech.

Almost all of the display settings on the available binocular head-mounted displays for testing are fixed and are not adjustable. Before comparisons in performance of these head-mounted displays for use in training can be made certain aspects of the

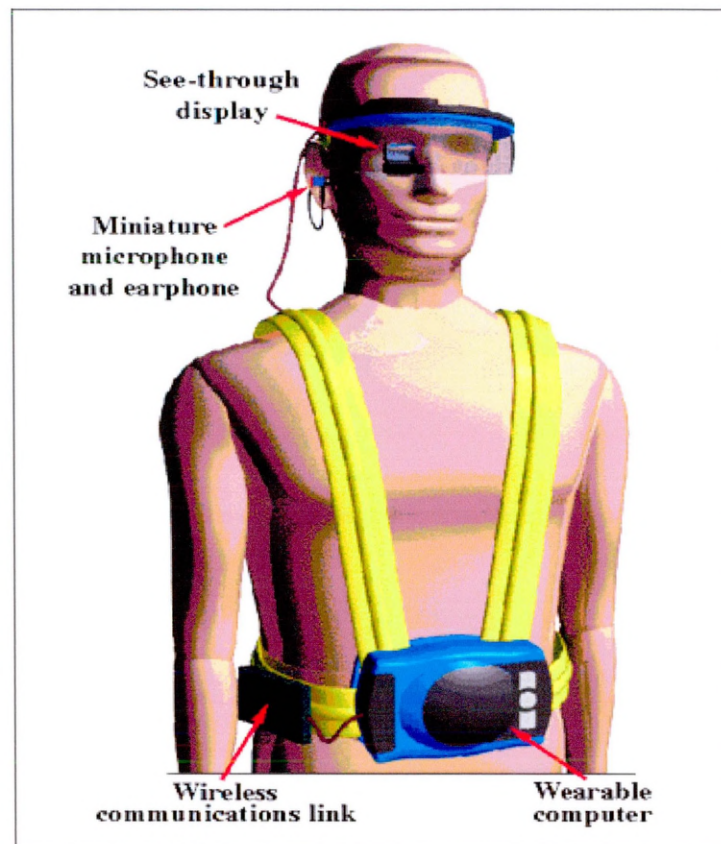
hardware would have to be identical. The field of view, distance of projected image and the resolution of the image itself would have to be equal. The only monocular head mounted display available to test was the “Liteye-D” and this had to be fixed on a hat or headband (See Figure 4.3). The display on this system produced a monochrome image that is not comparable to the image produced by the binocular head-mounted displays used in this research. After these problems became apparent inquiries were made about purchasing other head-mounted displays that produced high quality colour VGA images. However there were no other head-mounted displays that were reasonably priced and fitted the criteria needed to resume comparison studies of binocular and monocular display sub-types. Most other research projects have custom built wearable computers for the purpose of research. Ockerman et al. (1996) describe F.A.S.T. (Factory Automation and support technology), an example of such a specially created system that was developed at Georgia Institute of Technology (See figure 4.4). This F.A.S.T. wearable computer system was custom built for three types of information support, reference, information, task specific training and expert advice. To test the efficiency of different types of head-mounted displays for a specific task requires the creation of a custom built head-mounted display that can recreate the screen positions of monocular and binocular displays with all the visual configurations held constant. Given the incompatibility problems between the different screen images produced on different types of head-mounted displays it would be difficult to determine whether a significant difference in the performance of a head-mounted display was due to a

Figure 4.3: Liteye-D Amber Greyscale Microdisplay (Adapted from user manual 1999 and www.LYTEYE.com)



Using a 600 X 480 miniature, high resolution, AMEL display, the Liteye-D produces images, for viewing computer graphics and video images from a FLIR (forward looking infra red) unit. The display has a 35 degree field-of-view. The Liteye-D accepts both VGA and RS-170 Video signals.

Figure 4.4: F.A.S.T. Wearable Computer System. Georgia Institute of Technology.



The FAST hardware consists of a wearable, voice-activated computer system. This computer system was assembled in a lab and is equivalent to a 486 25Mhz desktop computer. This computer system consists of several components. The components are described below.

A see-through display allows the user to work while looking at text, drawings, and video that are pertinent to the user's job.

A miniature microphone/earphone provides audio information to the user and accepts voice input from the user that keeps the user's hands free for job-related tasks.

A wireless communications link sends and receives up-to-date information to and from the plant computer system.

A battery pack to supply power for all the components.
Adapted from (Ockerman et al., 1996)

difference in the projected image, field of view, position of head-mounted display screen or variation of the media in the instructions for the training programme.

Ellis & Menges (1998) provide evidence that the age of the user is another factor that may affect performance on different types of head-mounted displays. The study compared monocular, bi-ocular and stereoscopic displays for presenting an object. The dependent variable was the judgement of the distance of a floating triangle shape from the viewer. The independent variables were participants' age, accommodation (required focus) and the position of the physical surface. One of the results from this study showed that for all participants a monocular head-mounted display had a poorer performance than non-monocular headsets at optical infinity. At a nearer distance of fifty centimetres younger participants were able to judge distance better than older participants in the monocular display condition. A person's eyesight changes through time, they may experience a decrease in the ability to accommodate as they get older leading to a condition known as presbyopia (Goldstein, 1999). This raises issues about the eyesight of older participants using monocular systems and the need for focus control for all users. An important consequence of individual differences in eyesight is that all participants may require sitting a lengthy battery of eye tests before they take part in studies using head-mounted displays.

4.2 Evolving design of head-mounted displays and wearable computers.

The lack of adjustability to allow for individual differences in human vision is one factor that will affect comparison between different head-mounted displays. Another important aspect of head-mounted display technology is that the design of the headsets is a constantly evolving process. According to Yanagihara, Kakizaki,

Arakawa, & Isoda (1998) interaction between humans and wearable devices will become more sophisticated through advances in computer technology and more compact wireless communication devices. One of the head-mounted displays used in the first phase of the experiments is no longer being manufactured. This headset, the Sony Glasstron was designed and put on the market in the late nineties. Whereas the design of this head-mounted display may or may not be ephemeral, technology in successive head-mounted displays however, will be more sophisticated and advanced. Typical of the new twenty first century designs is the “spectacle-type wearable design described by (Tomono, 2000). This design constitutes a new generation of head-mounted displays where the spectacles themselves contain all the optical components for a see-through head-mounted computer display unit. This unit comprises of several components including the light source and lens incorporated in the spectacle glasses. This system has enough resolution to display SXGA level (1280 X 1024). This spectacle-type design is indicative of attempts by optical engineers to simplify the optical system of conventional head-mounted displays by projecting the image directly onto the retina.

A further important dimension of wearable computer systems is the method of interaction between the user and the system. Just as there are different types of head-mounted displays in wearable computer systems, there are various interaction devices within these systems that act as replacements for the mouse and keyboard. Bass et al. (1997) relate that these input devices come in many forms and interact with the system in different ways; they include dials, microphones for voice control, tracker-balls or three button input devices. These devices can be held in the hand or mounted on the body and normally come under the rubric of “keyboard surrogates”. If research

were undertaken to ascertain the optimum combination of head-mounted display and multimedia presentation to perform a task, then this superlative combination would have to be judged in the context of a compatible interaction device that would facilitate usability for a particular task. This may mean that the combination of a head-mounted display and a particular multimedia configuration alone does not provide enough information about a particular headset's performance in a wearable computer system. Whole wearable systems may have to be tested against one another for effectiveness in executing tasks.

Given the existence of the various factors that may potentially confound comparisons of head-mounted display subtypes, a decision was made to focus the remainder of the experiments in the research programme on ascertaining the optimum configuration of multimedia for the learning and retention of a procedural task. To continue the comparison of technologically different head-mounted display subtypes for performing procedural tasks would orientate the research towards more optical issues and away from investigating the configuration of multimedia in computer-based emerging technology. The design of wearable computing and mobile technology may change but the instructional dynamic visual display will always remain an important element of the instructional materials used in emerging technologies. The intention of the second phase was to use one particular head-mounted display type for viewing the manipulated multimedia instructions on a dynamic visual display. This would entail investigating the optimum method of placing sound and text on video instructions.

4.3 Chapter summary

- This chapter describes the problems encountered in the first phase of experiments regarding comparisons between head-mounted subtypes for following procedural instructions. Comparisons are difficult due to individual differences in visual characteristics of the users. The lack of adjustability of the subtypes of head-mounted displays makes it difficult to allow for individual differences such as interpupillary distance.
- Technological differences between the head-mounted displays also make comparisons problematic. The display screens in head-mounted display that receive a video signal are much harder to read than headsets that receive SVGA signals. Viewing information on the latter would be easier whereas instructions especially text may seem more distance and blurred on the former.
- Other technological differences between the head-mounted display subtypes are field of view and projection distance of the images. These two properties are often different in the subtypes and there is no opportunity to adjust these properties in order to keep these characteristics constant when testing for differences made by such issues as the position of the screen in the headsets. It becomes difficult to ascertain differences between the subtypes have affected performance on a task.

- The second phase of experiments focused on the optimum configuration of text, sound and visual information presented on a video for the learning and retention of a procedural task. The reason for this shift was due to the fact that further comparisons between head-mounted displays made available for study in this research programme were not viable and that an optimum design could not be identified through experimental evaluation.

Chapter 5

5 Introduction to second phase of experiments.

“I have always believed that the members of this house should be sufficiently articulate to express what they want to say without diagrams.”

The speaker Betty Boothroyd after rebuking an M.P. for using a diagram to explain overseas trade figures. (Guardian, 7 December 1994) Cited from Scaife & Rodgers (1996).

5.1 Rationale behind second phase of experiments.

The second phase of experiments attempted to address the fourth research question outlined in chapter one of the thesis (see section 1.2). The second phase investigated the best presentation of vocal and text instructions on a video clip for learning a procedural task. The main purpose of this investigation was to determine whether current guidelines for text and sound on a dynamic visual display could be used to design multimedia demonstrations for a procedural task. This chapter first looks at the evidence that procedural and declarative memory are two separate systems that encode visual and verbal information in separate ways thus may require different presentations of multimedia for learning. This chapter then goes on to describe the popular and widely used guidelines mentioned above that were created through years at the University of California Santa Barbara for constructing dynamic visual displays for learning scientific explanations. This section will go on to explain that these guidelines are based on research into declarative knowledge. There will also be a description of the three major influences on these guidelines for multimedia, dual

coding theory, cognitive load theory and constructivism. The chapter will then describe the three elements of multimedia that the second phase of experiments investigate; the use of video technology, sound and text in computer-based multimedia packages.

5.2 The architecture of human memory.

Michas & Berry (2000) put forward an argument through experimentation that guidelines for multimedia learning based on declarative research may not be appropriate for procedural learning. This argument is based on the concept that procedural memory and declarative memory are two distinct and separate systems. Guidelines based on declarative learning may not be applicable to the creation of multimedia learning materials for procedural tasks if these memories are encoded differently and stored in different parts of the brain. This section gives an overview of the evidence for this dichotomy, briefly describes the current theories about the architecture of human memory and exams the evidence from normal and abnormal patients for two separate memory systems.

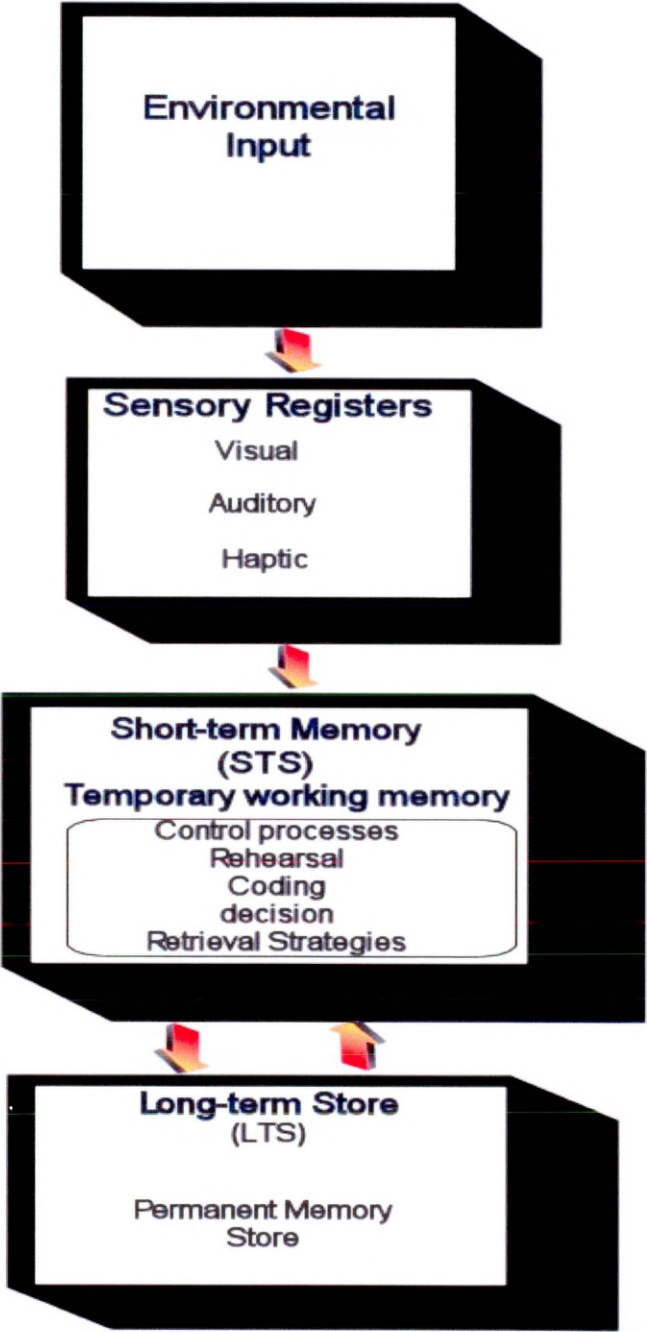
In recent years explanations of memory based on cognitive psychology have dominated the memory research field. In cognitive psychology learning is regarded as the development of the connections between a system of cognitive structures such as concepts, processes, facts and other types of information. New information is connected to previous information thus becomes easier to retrieve. Cognitive theories of learning and memory involve a series of stages; the first stage is a process known as encoding. Baddeley (1995) describes encoding as the initial processing of the information that is to be learned or memorized. Encoding results in some of the

information being stored in the memory system, this constitutes the second stage in the memory system. The third process in this memory system is the retrieval stage when the stored information is accessed again.

Parkin (1993) explains that in recent years the multi-store model has dominated memory theory. The increase in the use of computers in the 1960s led psychologists to draw similarities between computer processing and human memory function. The computer is viewed as a large database operated on by a central processing unit. This unit symbolizes a workspace where new data enters, existing information is retrieved and decisions involving the database executed. This set up is similar to an older concept of human memory that envisaged two memory types; short-lived memories that came under the label “primary memory” and longer lasting or permanent memories that were categorized as “secondary memory”(James, 1890). In the multi-store model primary memory represented the computer’s central processing unit and secondary memory was similar to the computer’s database of stored information. This concept of different types of memory led to a plethora of “store models” the most influential of which was the Atkinson and Shiffrin multi-store model (as cited in Baddeley 1987) (see figure 5.1).

The diverse theoretical models had similar characteristics, in most of these models, three types of memory store were envisaged. Sensory stores which hold information very briefly and are modality specific, a short-term memory of limited capacity and a long- term store of unlimited capacity that in theory holds information indefinitely. The multi- store model depicts memory as entailing the flow of information between

Figure 5. 1: Atkinson & Shiffrin (1968) multi-store model. Adapted from Baddeley (1987).



the three memory stores. New information first goes into the “sensory store” which is a very transitory store that holds information about the configuration of sensory information. Sensory storage of visual information is referred to as iconic memory and its auditory equivalent is known as echoic memory. Information in the sensory stores then passes into the short-term store. This store represents the seat of conscious mental activity and the actions of the short-term store symbolize the diverse control processes that mediate information passing into the short-term store. The control processes determine the contents of the short-term store in that the information being processed can be replaced by new information. According to the multi-store approach, Information in the short-term store that is subject to rehearsal transfers from the short-term store to storage in long-term store.

5.2.1 Long-term memory store.

Baddeley (1997) relates that experiments carried out in the sixties produced evidence that supported a short-term and long-term memory storage dichotomy and supported the computer analogy of memory processing. Experiments into the free recall of word lists for example demonstrated the recency effect (Glanzer & Cunitz, 1966). This involves a tendency for the last few items in the list of words to be initially very well recalled but after a short delay this recency effect disappears. One explanation of this phenomenon is that the more recent items are still being held in a transitory and delicate short-term store whilst earlier items from the word list are being recalled from the long-term store. Waugh (1970) discovered that items remembered from the recency part of a list of words were recalled faster than for the responses for earlier items. The above research collectively produced evidence for a dual memory system

with a short-term store with limited capacity but fast retrieval and a long-term store with unlimited storage but slower recall.

5.3 Two subtypes of long term memory.

Tranel & Damasio (1995) relate that a great deal of information about the neural foundations of human memory came from patients with rare patterns of brain damage. Amnesia has made a considerable contribution to the understanding of normal memory function (Parkin, 2001). A sizable amount of information about normal memory function was gleaned from a single patient called H.M. This patient developed anterograde amnesia after a bilateral resection of the medial temporal lobes including the hippocampus and amygdala to control epileptic seizures. After the operation this patient could no longer form new factual or declarative memories. Scoville & Milner (1957) studied the memory deficits of H.M. and these researchers came to several important conclusions about the function of human memory. First, the capacity to obtain new memories is a distinctive cerebral function that is localized to the medial part of the temporal lobes, independent from other perceptual and cognitive abilities. Second, the medial temporal lobes are not needed for immediate memory since H.M. was shown to have adequate short-term memory. H.M. could retain a visual image or a number for a short period time. Another conclusion from this research was that the temporal lobe couldn't be the sole storage area for long-term memory since H.M. still retained memories from his childhood.

H.M. could not remember new information such as names and faces, stories, drawings, odours, objects or music. It did not matter if the new information was

received aurally or read. H.M. appeared to hold the information in short-term memory however this memory could not be converted into long-term memories. Another remarkable finding from studying the memory function of H.M. was that he was unimpaired in learning a mirror drawing task. H.M. learned successfully to trace between the two outlines of a star while looking at his hand in a mirror. He exhibited daily improvement in this task, yet on each day he had no recall of having done the task before. This was taken as evidence that this type of procedural memory task is not dependent on the medial temporal lobe (Squire & Kandel, 1999).

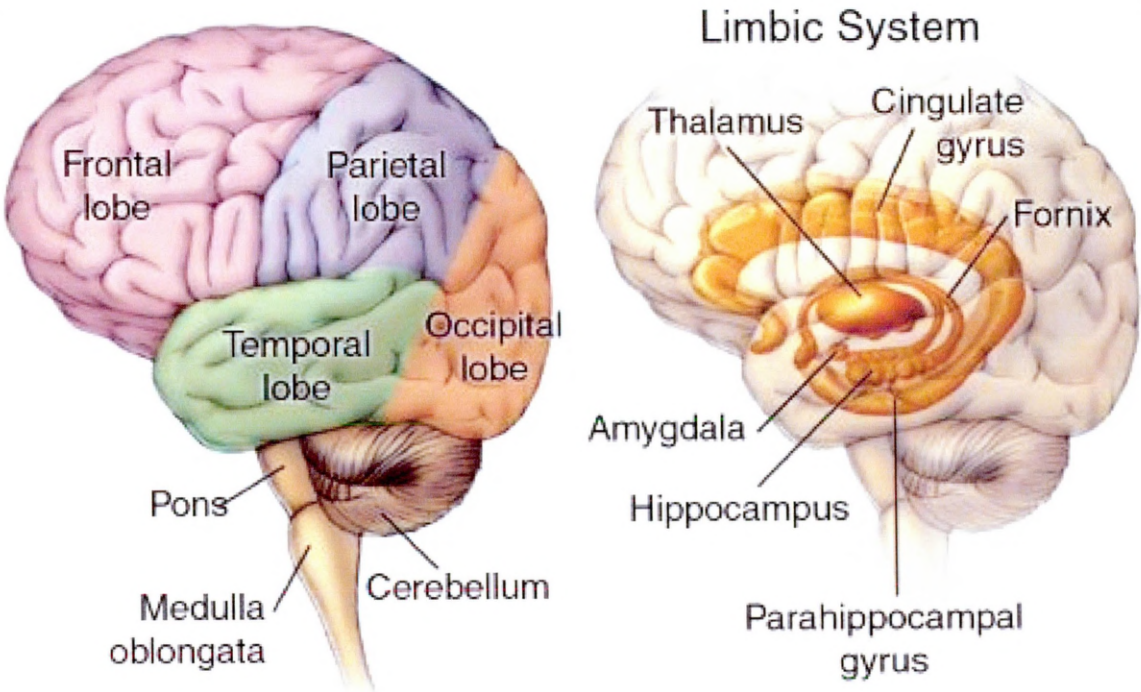
Squire & Cohen (1980) produced similar evidence from other patients with amnesia and concluded that long-term memory may also comprise of sub-systems by making a critical division between knowing how (procedural memory) and knowing what (declarative memory). Procedural learning is characterized by acquiring perceptual-motor and pattern-analysing skills belonging to a class of operations controlled by rules or procedures that cannot be easily recalled. Declarative memory incorporates the acquisition, retention and retrieval of knowledge that can be intentionally recalled. However there may not always be a sharp dichotomy between the two classifications for all tasks. Amnesic patients in the (Squire & Cohen 1980) study were able to learn mirror reading, a procedural task at a similar speed as normal participants and were able to preserve this ability for three months. It was concluded that amnesia appears to spare information based on rules, steps or procedures but impairs information that is declarative or fact based.

Some theorists of memory function have hypothesized that the mechanics of encoding new material would differ in declarative and procedural memory structures. For example, Eichenbaum (1997) postulates that declarative memory entails the encoding of memories in relation to the pertinent associations between the items. A fundamental characteristic of this form of memory is its ability to be flexible in its representations, an attribute that allows the inferential use of memories in novel circumstances. In contrast to this, procedural memories entail single representations; these memories are remote in that they are encoded only in the brain areas in which perceptual or motor processing is occupied during the learning of a skill. These single representations can only be exposed through the reactivation of these areas within the limited scope of stimuli and situations in which the original learning occurred, this renders these memories largely inflexible.

5.3.1 Evidence for the locus of declarative and the locus of procedural memory.

Squire & Zola-Morgan (1991) identified the structures in the medial temporal lobe involved with declarative memory for facts and events. Based on various studies of humans and animals it was postulated that amnesia was produced by damage to the hippocampus and amygdala, which reside in a brain area known as the limbic system (see figure 5.2). Since declarative knowledge acquired before the lesions or brain damage was still accessible, it was posited that the structures in the medial temporal lobe may only consolidate the memories by gradually binding together information from the various cortical areas that store a whole memory episode before moving them to a new permanent store, possibly in the neocortex (Squire & Zola-Morgan, 1996).

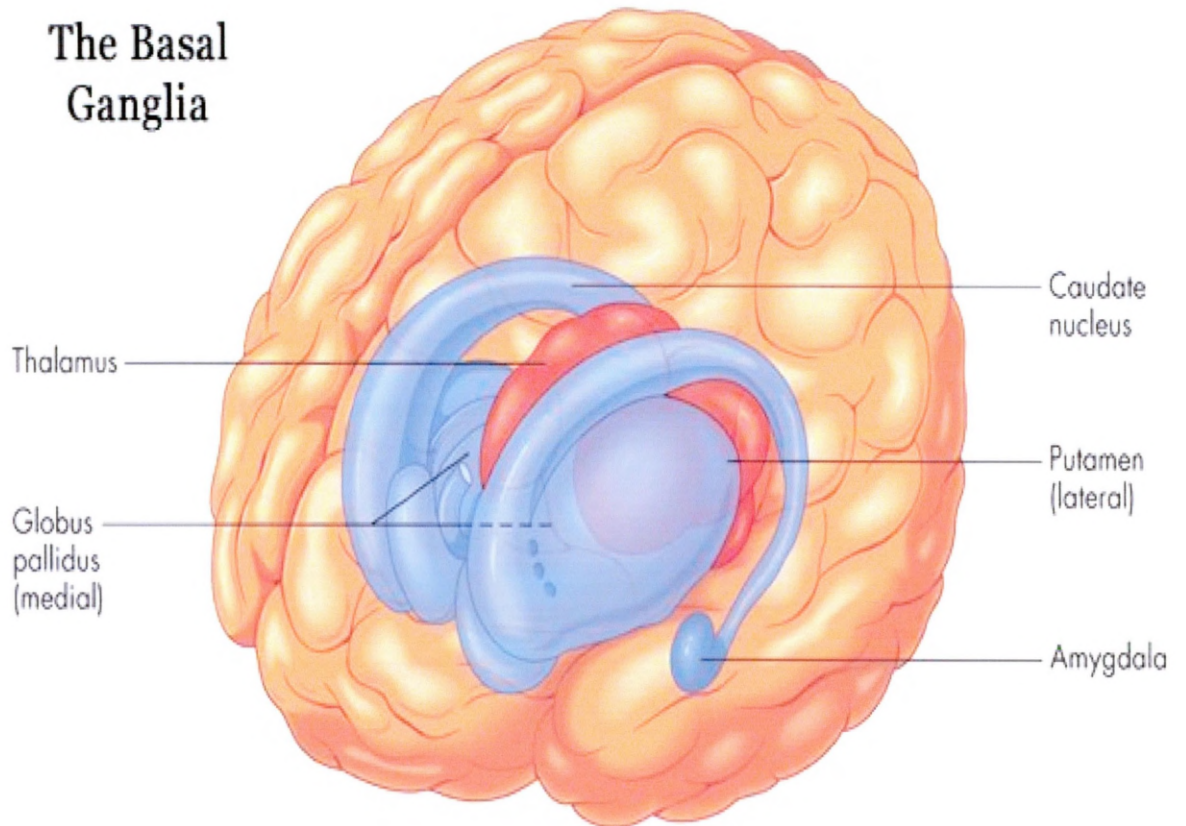
Figure 5.2: Brain regions associated with declarative memory. Adapted from www.ahof.org/aldis/about/AnatomyBrain.htm



Whereas evidence from amnesiacs has shown deficits for declarative knowledge, data from several studies of memory impairments of Huntington disease and Parkinson disease patients have demonstrated procedural memory deficits for these patients. As explained by (Butters, Wolfe, Martone, Granholm, & Cermak, 1985) Huntington's disease is a genetically transmitted syndrome that results in the gradual atrophy of the basal ganglia, specifically the caudate nucleus. The symptoms of this disease are characterised by jerky limb movements and a progressive dementia, problem solving and memory deficits. Memory impairments in Huntington's disease have been cited as clear evidence for multiple memory and learning systems in the human brain. Patients with Huntington's disease have been shown to be poor at learning skill based information and procedural tasks. Parkinson's disease has a characteristic aetiology that includes significant degeneration and atrophy of the basal ganglia, particularly the caudate nucleus (see figure 5.3).

Research into memory deficits of Parkinson's disease has contributed to the debate in dissociating different classifications of memory and the areas of the brain that mediate different memory or learning processes (Hay, Moscovitch, & Levine, 2002). In one study, Hay et al. (2002) used a habit-learning task was to compare Parkinson's disease patients to amnesics and controls for procedural memory. The results from this study showed that the amnesic patients exhibited the normal dissociation of impaired recollection and spared skill or habit learning. Parkinson's disease patients were impaired in their ability for habit formation and their ability to recollect specific information. The conclusions from this study was that Parkinson's disease is characterised by a deficit in learning automatic motor skills due to striatal

Figure 5.3: Brain areas associated with procedural memory. (Adapted from www.bvu.edu/faculty/ferguson/biotech/images/basalganglia)



degeneration whilst frontal lobe dysfunction may contribute to the deficits in recalling information

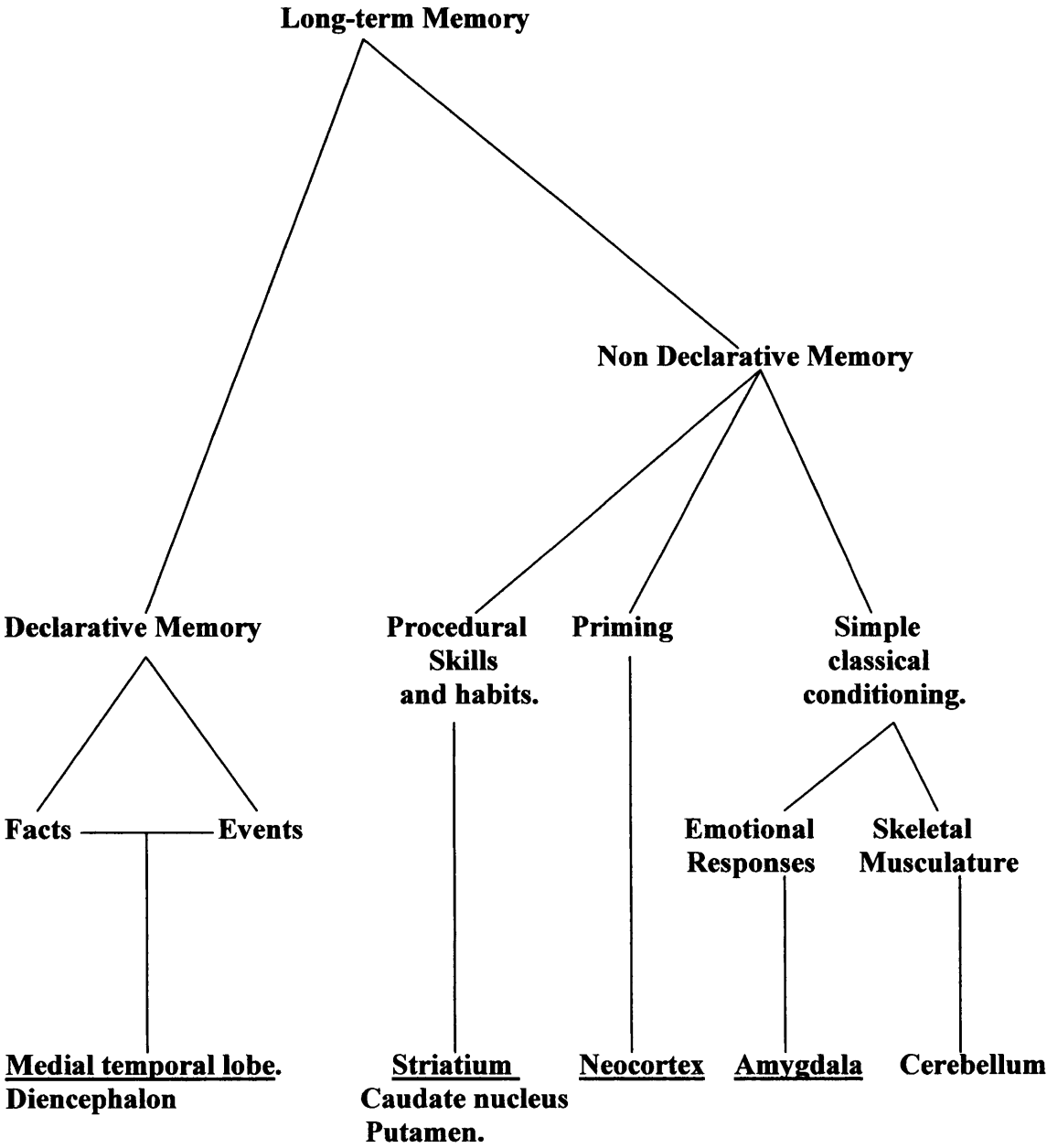
Another area in the brain associated with procedural memory is the cerebellum, which sits beneath the cerebral cortex and behind the brain stem. Fabbro (2000) report that cerebellar lesions produce various neurological symptoms. Studies of these impairments has allowed researchers to speculate that the cerebellum is involved in a variety of functions including regulation of linguistic, cognitive and emotional processes as well as learning and procedural memory, visuo-spatial tasks, attention and sensi-motor tasks, language and verbal memory. One of the main functions of the cerebellum is to establish associations, to link the context in which the movement is made to the lower level movement generators. The cerebellum is involved in learning and making complicated motor sequences automatic. The collective research from brain damaged and diseased patients produced a taxonomy of proposed long-term memory systems along with the particular brain structures implicated in each system (see figure 5.4).

5.3.2 Implicit memory and multiple memory systems.

Parkin (2001) records that human memory research in the 1990s became influenced by a new trend, the investigation of implicit memory and implicit learning and this had an impact on theories about the nature of procedural and declarative memory.

Berry & Dienes (1991) describe two research areas that became prominent in cognitive psychology: the implicit versus explicit memory dichotomy and the implicit versus explicit learning dichotomy. The explicit and implicit memory approach tries to comprehend long-term memory by exploring how it reacts to implicit and explicit

Figure 5. 4: Schematic illustrating systems and brain structures involved in long-term memory. (Adapted from Squire & Zola-Morgan (1996)).



memory tests. Implicit memory tests are defined as those, in which memories for past experiences occur without needing conscious access to these memories. On the other hand, explicit memory tests require the conscious recall of past events.

The second research area that became influential was the differentiation between implicit and explicit learning. Smith, Siegert, & Mcdowall. (2001) used a common implicit learning test, the artificial grammar task to test the existence of implicit learning in Parkinson's disease patients. The results of this study demonstrated that Parkinson disease patients could learn an artificial grammar task. Artificial grammar learning involves the presentation of a set of rule-governed stimuli to participants. The stimuli are usually letter strings from a finite-state grammar that are studied by the participants. The set of stimuli usually comprises of exemplars that are representations of the complete range of transitions of the grammar. These transitions give exposure to the rules of the grammar in its entirety in an indirect way. After the presentation of the examples of the transitions the participants are notified about a complicated grammar system mediating the presented stimuli. Participants are then shown new word strings only half of which conform to the grammar rules and the participants must decide which word string corresponds to the grammar structure. The underlying postulation of this experimental paradigm is that tacit knowledge, which is abstract and represents a complex grammar system, can be learned unconsciously. The distinguishing features of implicit learning and implicit memory research are that the former entails the study of the acquisition, retention and retrieval processes whereas the latter is only concerns the retention phase of the memory system.

Reber (1993) voices the opinion that the two research programmes implicit memory and implicit learning have for the last twenty years progressed in parallel with little interaction. Some researchers, for example (Berry & Dienes, 1991) have tried to forge links between the two areas of study by distinguishing common characteristics. One of these is the convergence of between theoretical descriptions of the two phenomena. Implicit memory and implicit learning have been placed in the framework of multiple memory systems. Theorists, for example (Squire, Knowlton, & Musen, 1993) have taken explicit memory and implicit memory as evidence for the two distinct memory structures. Explicit memory tasks are seen as declarative in nature, but implicit memory is seen as procedural.

5.4 Guidelines for multimedia learning.

Phase two in the experimental programme sought to investigate the fourth research question in the thesis, namely the optimum configuration of multimedia information on an instructional video clip for learning a procedural task. Weiss, Knowlton, & Morrison (2002) comment that although there exists a substantial body of research into animations and dynamic moving displays, a limited amount of this research has produced guidelines or heuristics for the use of these moving displays in multimedia environments. Only a handful of researchers have engaged in research work that has striven to create guidelines for the creation of multimedia learning materials. As far as the construction of multimedia learning materials and the placement of text and verbal information in animations is concerned, the theories and principles produced by Mayer and co-workers derived from their empirical research have become regarded as mainstream in multimedia design.

For over ten years Mayer and his research collaborators at the University of California Santa Barbara conducted dozens of studies on multimedia learning. Their research programme started with pictures and text and progressed to animations with voice narrations. According to Mayer (1989) the starting point for conveying learning through multimedia is that the material to be learned should be potentially meaningful, in other words there should be a propensity for creating a coherent mental model from the learning materials. The type of declarative learning that Mayer and colleagues focused on for these studies was the understanding of scientific text, specifically cause and effect explanations of how systems work.

The scientific explanations Mayer and colleagues used in their studies were: an explanation of how lightning works (Mayer & Moreno, 1998), an explanation of how a bicycle-pump works (Mayer & Anderson, 1992) and an explanation of how a brake system works (Mayer & Gallini, 1990). This body of research work, which was mainly committed to discovering the optimum method of presenting visual and verbal information in multimedia animations and presentations culminated in the creation of a “Cognitive theory of multimedia learning” (Mayer, Bove Bryman, Mars, & Tapangco 1996; Mayer, Heiser, & Lonn, 2001). This theory is derived from three theoretical frameworks: dual coding theory, cognitive load theory and constructivist learning theory. Taken from the dual coding theory is the idea that visual and verbal materials are processed in different cognitive systems. Borrowed from the cognitive load theory is the idea of limited processing in the verbal and visual channels. Finally, the model has been influenced by the concept in constructivist theory that meaningful learning occurs when learners select appropriate information and integrate it with

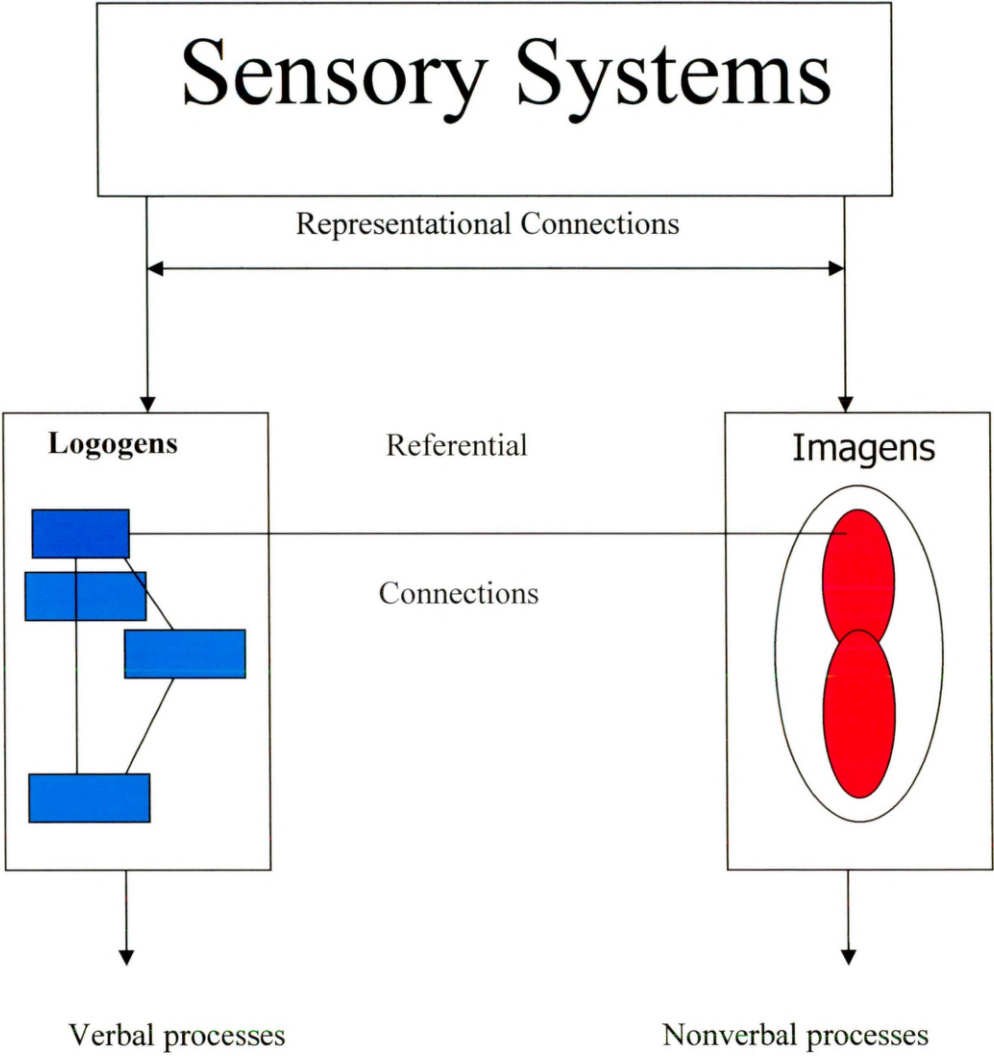
previous knowledge. These theoretical perspectives have all influenced the cognitive theory of multimedia learning.

5.5 Dual coding theory.

Molitor, Ballstaedt, & Mandle (1989) recount that the “dual coding” theory (Paivio, 1986); (Paivio, 1991) is the most common and influential theoretical approach to understanding the cognitive basis of multimedia. This theory has focused on the memory storage of verbal and visual information and postulates that the combination of visual and verbal learning materials has an additive effect on learning. This concept has greatly influenced the research work of Mayer and colleagues and much of their theoretical models are derived from the dual coding theory. This dual coding theory assumes that two distinct processing systems exist, one for verbal information and one for visual information. These systems function independently but interact with one another. Not only are these processing systems functionally distinct but also are structurally dissimilar since they store information in modality specific representational units called logogens and imagens (see figure 5.5). The logogens symbolize verbal codes such as text or speech and the imagens represent natural objects and are stored as mental images of the properties of these objects.

A second structural difference between these two representations is the manner in which they are organised. The imagens have the benefit of being structured in a simultaneous way that permits several elements of an image to be processed synchronously. In contrast to this logogens are organised in bigger units and in a serial manner, therefore susceptible to the confines of sequential processing, which

Figure 5.5:Schematic of Paivio’s Dual Coding Theory (adapted from Paivio 1986).



limits the amount of information processed at a time. Although presented as a dichotomy of two functionally separate cognitive systems, these sensory systems are interconnected. The visual and verbal representations can form associative connections that allow the conversion of each type of information into the other. An example of this would be that reading the word bird would be instantly connected with a mental picture of a bird and hearing the word bird might also create a mental picture of a bird (Paivio, 1986; Paivio, 1991). Clark & Paivio (1991) postulate that the dual coding theory has important consequences for education. Multi-modal teaching materials may enhance the efficiency of instruction by providing learners with the possibility of storing the same information in both a verbal and a visual memory representation. The dual coding theory proposes that when the same information is presented simultaneously in a verbal and a visual format, learners form connections between the visual and verbal information during encoding. The advantage of this dual coding is that it may augment the number of routes that learners can use to retrieve information since verbal stimuli might initiate verbal and visual representations. Learning materials that comprise of both text and pictures might aid retention since they furnish learners with two methods of memorizing the materials.

Another important implication of the dual coding theory is that concrete information is more likely to be remembered than abstract information using words and pictures. Paivio, Clarke, & Khan (1988) contend that concrete information is better remembered due to the fact that it can produce mental images and aid people to encode the same learning materials in the two modalities. A concrete word is a coded as a word and a visual image whereas an abstract word is coded solely as a word.

5.5.1 Research involving pictures and text.

There has been a large volume of research into the dual coding theory that used picture and text as the media in the studies. In the last thirty years Paivio and colleagues have carried out studies to investigate participants' recall of visual and verbal information (Paivio & Caspo, 1973; Paivio, Walsh, & Bons, 1994). In one experiment Paivio & Caspo (1973) asked participants to memorize lists of either words or sentences and pictures showing concrete ideas and asked them to remember them later. A persistent result from these studies was that there was a superior memory for pictures than words (Paivio, 1983). A secondary finding was that processing both words and pictures of the same material had a cumulative effect on memory. In these studies participants who were in condition where the material to be learned was presented both as text and pictures recalled more words than those who were presented the information as text or pictures alone (Paivio, 1983; Paivio & Caspo, 1973).

Educators and researchers who were interested in the manipulation of media to improve learning were heavily influenced by the dual coding theory for retaining concrete information. A number of researchers in the seventies and eighties attempted to investigate the dual coding theory in an educational environment. These studies were carried out in a variety of learning situations with different learning materials. In a substantial review of the effects of text with illustrations compared to text alone, (Levie & Lentz, 1982) reviewed 155 studies that used 7182 participants. Students of a variety of ages were used. For the youngest participants children's stories were used with line drawings and for older students explanatory texts were used in the studies along with pictures or diagrams. The main conclusions drawn from this meta analysis

was that illustrations in general support learning and enhance learning from text, although it was not certain from these studies how this mechanism works. Some of this research indicated that illustrations drew attention to non-compulsory reading, however there was no hard evidence that the illustrations had an effect on the attention of textual information that participants were assigned to read.

5.6 Modifications of dual coding theory.

The researchers at the University of California Santa Barbara developed their own modifications of this theory to explain their research results in multimedia. The first modification of the dual coding theory (Paivio, 1986; Paivio, 1991) was called the “Integrated dual-code hypothesis (Mayer & Anderson, 1991). This model comprises of three component processes (see figure 5.6). This first component involves building representational connections between information presented verbally and a verbal representation. The second component entails creating referential connections between information presented visually and a visual representation.

The third component entails the creation of referential connections between elements in the verbal and visual representations. A second, later adaptation of the dual coding theory was called the “The dual-coding model of multimedia”(Mayer & Sims, 1994). This model was also a three-process account of how information presented visually and verbally may be integrated with working memory in the learning process (see figure 5.7). On the top left part of the figure a verbal explanation is presented to

Figure 5.6: The integrated dual-code hypothesis. (adapted from Mayer & Anderson 1991)

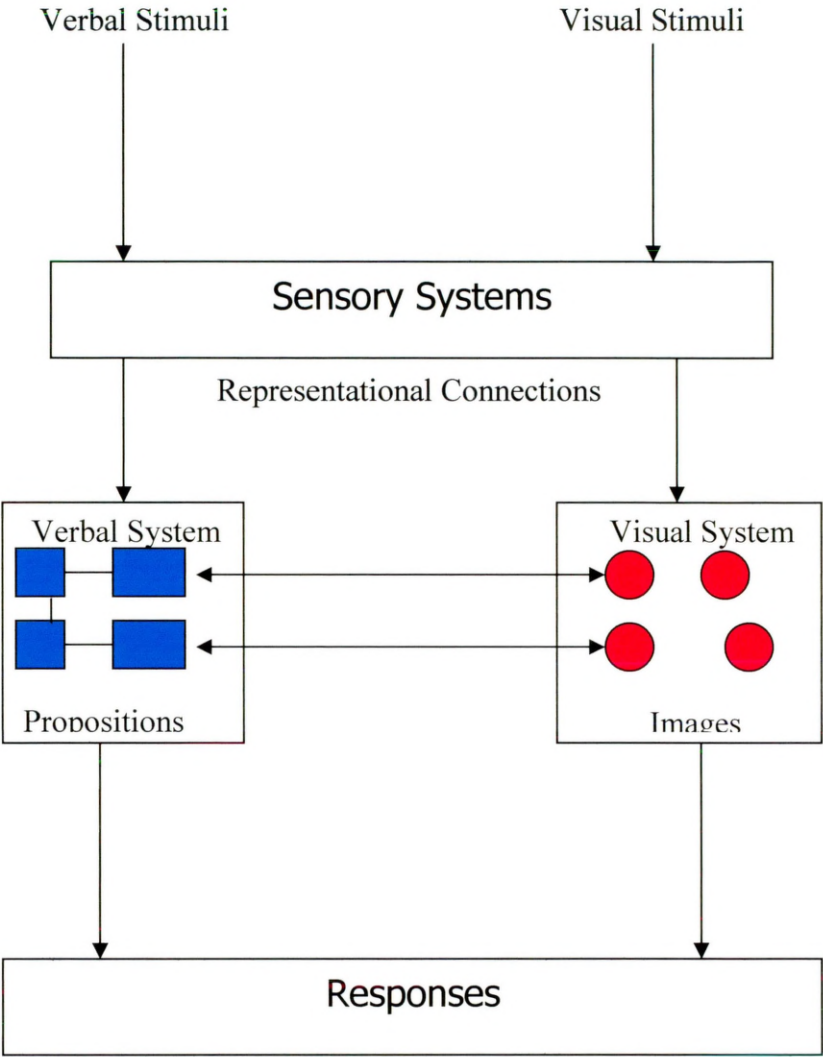
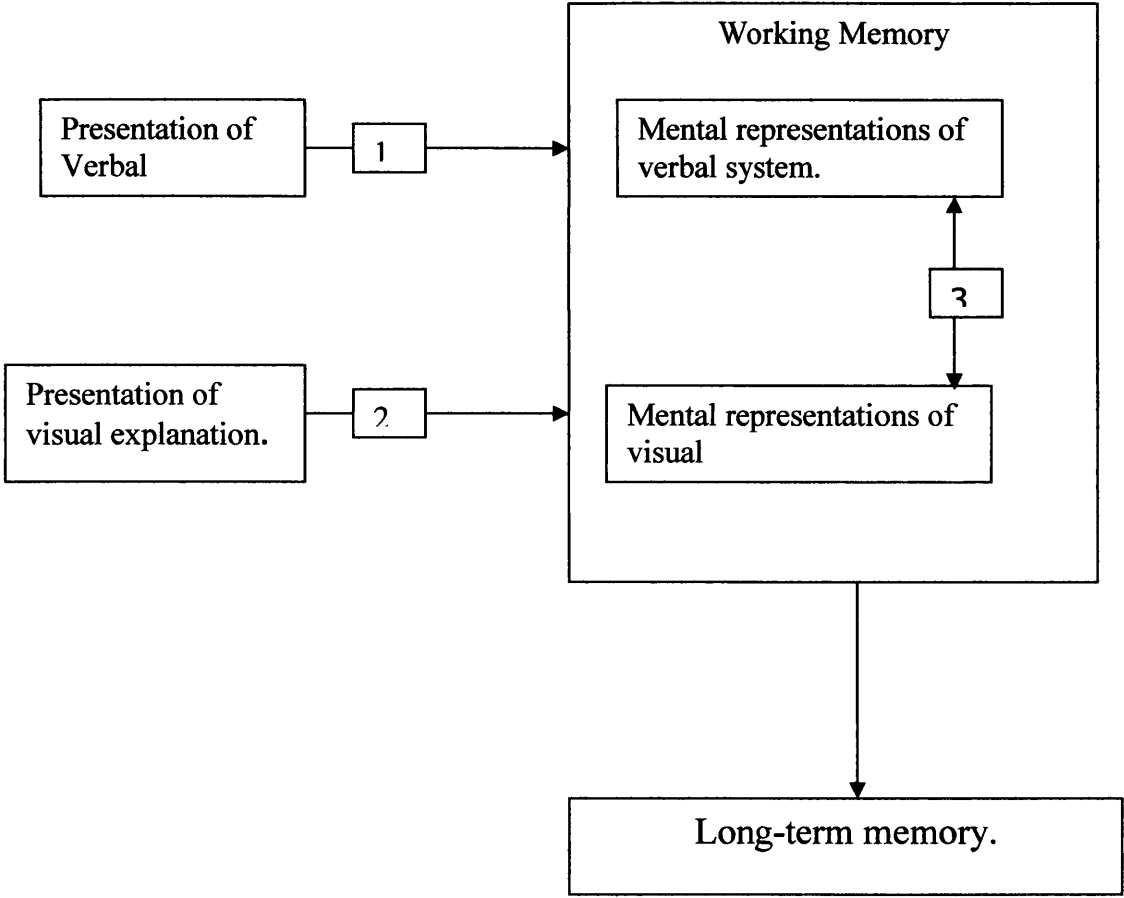


Figure 5.7: A dual coding model of multimedia learning. (adapted from Mayer & Sims 1994)



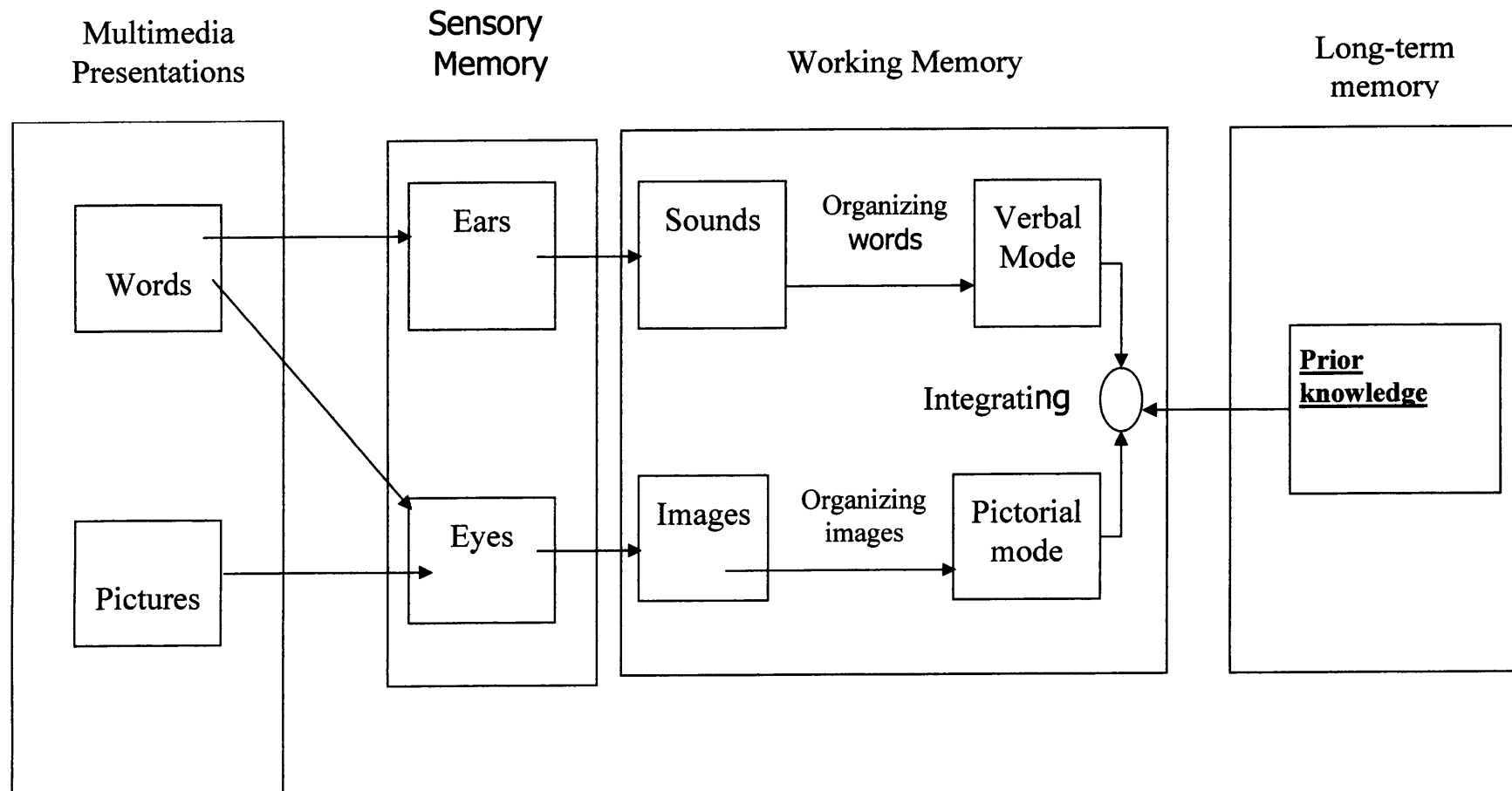
the learner. The learner then builds a mental representation of the system drawn from the verbal explanation. The cognitive process of moving from an external representation of visual information is called “building an internal representational connection” and is indicated by the second arrow in the diagram. The third arrow indicates the “construction of referential connections” between the two mental representations of the system.

Figure 5.8 depicts the cognitive theory of multimedia learning. This theory was influenced by the dual coding theory, cognitive load theory and constructivism. The top row symbolizes the visual channel. The boxes on the left of the figure represent the presentation modes for the multimedia instructional material, words are presented as narration and pictures are presented as animation. The learner attends to relevant parts of the animation (“select images” arrow) and holds these images in working memory (image base box). The learner will then mentally create connections that organise the words (“organise words”) arrow into a cause and effect sequence (“visual mental model” box). Finally the learner will construct referential associations between the visual and verbal models with prior knowledge (Mayer & Moreno, 2002a). This is a model of multimedia learning constructed to account for learning using animations coupled with narrations.

5.6.1 Cognitive load theory.

Another research paradigm that greatly influenced the method in which multimedia learning materials are created is the “Cognitive load theory” (Chandler & Sweller, 1991). Bannert (2002) describes the cognitive load theory as an internationally renowned and pervasive theory that has been supported empirically by many studies.

Figure 5.8: Cognitive theory of multimedia learning (adapted from Mayer & Moreno (2002))



The theory was developed over several years by Chandler, Sweller and colleagues. Cognitive load theory is based on cognitive theories of human architecture and the presumption that visual and verbal working memory has a limited capacity. The theory was evaluated empirically mainly by studies involving explanations of how electrical appliances work (Kalyuga, Chandler, & Sweller, 1998).

Not only has the cognitive load theory been instrumental in shaping ideas in multimedia instruction but it has also contributed to other learning environments such as the development of “computer supported collaborative learning” where external representations may be divided between students (Van Bruggen, Kirschner, & Jochems, 2002) . Kirschner (2002) credit the cognitive load theory with providing guidelines for the presentation of information in a way that stimulates learner interaction that leads to a maximization of learning outcomes. It is also valuable for creating training formats to satisfy the cognitive demands of older learners (Van Gerven, Paas, Van Merriënboer, & Schmidt, 2002). A major concern of the cognitive load theory is that the limited processing capacity of the human brain has restricted the acquisition of complicated cognitive skills. The theory is recognised for serving a major function in multimedia training by providing guidelines that address these cognitive limitations.

5.6.2 Theoretical framework of cognitive load theory.

Pollock, Chandler, & Sweller (2002) depict cognitive load theory as a theoretical framework that draws on certain elements of cognitive structure and the configuration of information to create designs for instructional materials that ease comprehension,

learning and problem solving. This theoretical framework is based on three assumptions. The first of these is that humans possess a limited short-term or working memory that can only process a few pieces of information at a time. Miller (1956) produced evidence that this short-term store is limited to approximately five pieces of information. More information than this limit may overstretch short-term or working memory, reducing the efficiency of processing. The second assumption that cognitive load theory is based on is that long-term memory in contrast to working memory has an unlimited storage capacity (Newell & Simon, 1972). The third assumption that forms the basis of this theory is that the information in long term memory is stored in the form of “schemas”, these are hierarchical frameworks which allow complicated knowledge structures to be treated as single elements in memory (Gick & Holyoak 1983; Thorndyke & Hayes-Roth, 1979).

Tindall-Ford, Chandler, & Sweller (1997) explain that these schemas are structured so that they store information in long-term memory in a way that circumvents the limitations of working memory. The last assumption cognitive load theory is based on is that there is an automation process that permits schemas to be processed automatically rather than consciously in working memory. This concept of automatic processing is based on work carried out by Schneider & Shiffrin, (1977) into human information processing. Automatic processing is the initiation of a learned sequence of actions in long-term memory that is started by specific inputs that trigger the automatic execution of a specific action sequence.

A fundamental assumption of cognitive load theory is that the presentation of the instructional materials should be organized in a manner that reduces extraneous

working memory load. One method of cognitive load reduction is the use of worked examples. Sweller (1988) hypothesized that novice learners solving a problem by a means-end procedure focused on elements of a problem that were not suitable for schema acquisition, therefore this method of problem solving is not an effective means of learning since it applies a heavy cognitive load. In contrast to this, worked examples have a greater propensity to develop efficient schemas since they focus the learners attention on problem states and operators rather than goals and sub goals (Van Gerven et al., 2002).

Another presentation effect that may mediate extraneous cognitive load is the “split attention effect”. This effect occurs when learners must focus on multiple sources of information that must be mentally integrated before meaning can be obtained from the learning material. The mental integration of text and visual information set apart results in a higher cognitive load and less effective acquisition of the information. Research has demonstrated that physically integrating text statements into diagrams rather than placing the text adjacent to the diagram circumvents the split attention effect (Chandler & Sweller, 1991). The final effect that has an influence on extraneous cognitive load is the modality effect. As noted above the physical integration of instructional forms alleviates the strain on limited working memory capacity. The modality effect occurs when multiple sources of information that need to be integrated are presented in two modalities. More information can be processed through two channels, auditory and visual than through a single channel (Mousavi, Low, & Sweller, 1995). The split attention effect and the modality effect have had a marked influence on the design of modern multimedia heuristics. These two effects in

cognitive theory heavily influenced Mayer and colleagues when they were devising experiments and creating the cognitive theory of multimedia learning.

5.6.3 Constructivism and computer based learning.

The second phase of experiments tested the presentation of multimedia instructions on a video clip in an effort to clarify some of the issues regarding the use of multimedia to enhance procedural learning. A majority of the principles for the design and development of multimedia in computer-based instruction mentioned above have been derived from constructivism. The three prominent theoretical frameworks of learning are behaviourism, cognitivism and constructivism, each of these three paradigms constitute a different approach to the design of multimedia instructional systems (Jonassen, 1991). The behaviourist theoretical perspective showed that the principles of operant conditioning could be successively applied to problem solving in the classroom environment. Skinner (1948) who was an influential proponent of behaviourism promoted the notion that learning is shaped by particular reinforcement by way of motivational or corrective feedback to augment the possibility of obtaining specific behaviours.

A criticism of behaviourism is that it cannot explain the complex nature of human learning (Kettanurak et al., 2001). The principles of behaviourism have been used in programmes involving low-level skills as opposed to learning complex conceptual behaviour. Some principles have been used successfully in the design of instructional packages, for example in the use of feedback as reinforcement, the breaking down of lessons into small fragments making programme choices based on visible learner

actions (Hannafin & Rieber, 1989). Behaviourism advanced knowledge about learning but behaviourist experimental findings from animal studies could not always be generalised to humans. As far as some instructional researchers are concerned, the behaviourist paradigm of learning has been eclipsed and superseded by ideas derived from cognitive psychology (Park & Hopkins, 1993). The cognitivist approach maintains that the learner processes information in a way that depends on an interaction between short and long term working capabilities. Learning is concerned with what learners know and how this learning is acquired. A number of guidelines for learning are derived from the cognitivist perspective, for example the progression from novice to expert is employed to correspond to the learner's mental state. However this perspective has been criticized for being limited for the construction and reconstruction of the learners' schemata for the changing environment (Tenenbaum, Naidu S., Jegede, & Austin, 2001).

The third theoretical perspective that has influenced computer-based instruction is constructivism, Jonassen (1991) defines the latter as a process where the learner constructs an inner representation of knowledge through an interaction with the environment. Constructivism explores the unprompted and self-paced learning typified by the natural and untaught method of forming ideas. Whilst the behaviourists posit that learning is the result of reinforced behaviour, constructivists argue that meaningful learning results only during personal reflection rather than through environmental factors. In recent years enthusiasm for constructivism has grown considerably at the expense of traditional learning methods which have been dismissed as being rigid and inflexible. Some educationalist, however, have attacked constructivism for having no structure and few restraints in contrast to the

methodologically planned instruction used in the behaviourist and cognitivist approaches.

Another common criticism is that the open-ended constructivist instructional materials may prove daunting to the novice learner (Kettanurak et al., 2001). On the other hand the constructivist paradigm has possible advantages for open and distance learning and other emerging technologies have incorporated constructivist applications into courseware (Tenenbaum et al., 2001). Winn (1997) outlines the advantage of using constructivist learning in virtual environments. Since the latter are rich and complex environments in which learners have the propensity to construct their own comprehension of learning materials in a similar way to the real environment. As far as general computer-based instruction is concerned, instructors have used the constructivist approach to create instructional programmes with tools such as databases, interactive multimedia with montages and expert information that can be used by learners to construct a knowledge base on a particular subject (Weeks, Lyne, Mosly, & Torrance, 2001). Learner control environments in multimedia instruction have been inspired by constructivist learning.

A study by (Mayer, Moreno, Boire and Vagge, 1999) outlined the importance of constructivist learning theory in the research work into multimedia demonstrations that took place at the University of Santa Barbara. Constructivist learning is important when learners employ an active learning construction process and select relevant phrases and images and organise them into the scientific explanations of how lightening or brakes work. The research work into using constructivist ideas for multimedia learning has however dealt with learning declarative information. One

question is whether constructivist ideas regarding the configuration of multimedia are applicable to learning procedural knowledge.

5.7 Principles for multimedia learning.

The researchers at the University of California Santa Barbara (UCSB) who undertook a body of research into the optimum method of presenting multimedia learning environments produced seven principles for the design of multimedia learning materials derived from the cognitive theory of multimedia learning. These principles are; the multimedia principle, spatial contiguity principle, temporal contiguity principle, the coherence principle, the modality principle, the redundancy principle and the individual differences principle. Mayer & Moreno (2002b) describe these seven principles; the first is the “multimedia principle (Mayer & Anderson, 1991; Mayer & Anderson, 1992) which holds that learners learn more profoundly from animation and narration together than from narration or animation alone. This principle is heavily influenced by the dual coding theory and is based on research that moved the ideas behind dual coding theory from pictures and text to animations. The theoretical rationale behind this principle is that when words and pictures are both presented together, learners have the opportunity to build connections between them.

The next principle is the “spatial contiguity principle” (Moreno & Mayer, 1999), which states that explanatory text should be integrated with pictures rather than separated from pictures. The theoretical rationale behind this theory is that when matching pictures and text are separated from one another learners expend greater cognitive resources to visually search the screen and are less able to hold the information in

working memory concurrently. The third principle is the “temporal contiguity principle” (Mayer & Sims, 1994), which is similar to the previous principle but states that learning is enhanced when corresponding elements of narrations and animations are presented at the same time. The learner has a greater chance of holding mental representations in working memory simultaneously. As well as this the learner has a greater propensity to construct associations between verbal and visual representations. This is a principle which is based on the split attention effect in cognitive load theory. This principle is an adaptation of the modality effect in cognitive load and again entails applying results from picture and text research to dynamic visual displays.

The fourth principle is the coherence principle (Mayer, 2001) when learners demonstrate improved learning when unrelated material is omitted from the learning materials, for example irrelevant pictures, sounds text and music. The fifth principle is the modality principle (Mayer & Moreno, 1998); this principle suggests a learning advantage for animation and narration rather than from animation and text. The theoretical basis of this principle is that when pictures and words are presented visually, as animation and text, the visual processing channel may become overloaded. When the verbal information is presented as narration however, this information can be processed in the auditory channel, thus allowing the visual channel to process just the pictorial information. This principle is derived from the modality effect in cognitive load theory.

The sixth principle in this series is the “redundancy principle”(Mayer et al., 2001). This principle holds that a multimedia presentation of animation and narration is better learned than a presentation consisting of animation, narration and on-screen

text. This is based on the concept that when words and pictures are both presented to the visual channel, this processing channel may become overloaded. The principle has its origins in the split attention effect from cognitive load theory. The final principle created from the research programme into multimedia learning was the “individual differences principle”(Mayer et al., 2001), this is where high-knowledge learners have an advantage over low-knowledge learners for using multimedia and high spatial learners have an advantage over low spatial learners for using multimedia learning materials. High-knowledge learners can use prior knowledge to make up for limited guidance in the multimedia presentation. High spatial learners have the cognitive capabilities to integrate mentally visual and verbal representations from successful multimedia presentations.

The principles and models for multimedia learning constructed by the researchers at UCSB have been based on teaching and conveying scientific cause and effect explanations and involve the learning of declarative and factual materials. The body of research that led to the creation of these principles had the dual coding theory as its starting point. This is a theory that is based on the processing of concrete material rather than abstract ideas or physical actions. These models and heuristics have been shown to hold true for other types of declarative learning but have not been empirically tested for learning different types of procedural tasks with dynamic visual displays. There is no body of similar research that has been based on theories for the acquisition of procedural tasks for example the “ACT” theory of procedural learning (Anderson, 1993). Despite this there has been a tendency in the multimedia learning literature to regard such heuristics and guidelines as effective for all learning whether declarative or procedural (Reimann, 2003).

The only criticisms of the theoretical frameworks regarding multimedia learning devised by Mayer and co-workers originate from propositional arguments against the importance of visual images in the multimedia learning process. For example, a study by (Schnotz & Bannert, 2003) used hypermedia and hypertext to evaluate their own model of integrated text and picture comprehension. This model was more concerned with role of visual images in creating a mental model of the information and the role of text in producing propositional representations. The contention put forward by this paper is that adding pictures to text is not always advantageous to learning. Only in certain circumstances is the combination of visual and verbal information beneficial; for learners with low prior knowledge and when the subject matter is imagined in a task appropriate way. The visualization in the images must support a task appropriate mental model. There has been no similar body of research into the construction of multimedia learning materials with verbal and visually presented instructions for a procedural task. Phase two will test four of the principles set out by Mayer and colleagues at the UCSB; the multimedia principle, the temporal contiguity principle, the, modality principle and the redundancy principle on multimedia instructions for a procedural task. These four principles are more pertinent to the design of multimedia dynamic visual displays than the other three.

5.8 The multimedia paradigm.

Hoogeveen (1997) terms the shift towards the use of multimedia in computer-based instruction as the “the multimedia paradigm” This movement is founded on the widespread belief that the addition of multimedia to an information delivery system

will inevitably lead to the improved learning and retention of the information.

Another common conviction about multimedia is that it is fun and exciting to use. As pointed out by (Winn, 2002) computer technology itself has a growing influence on education mainly due to the continuing reduction in price of powerful desktop computers, making them more affordable for schools and further educational establishments.

An important dimension of emerging computer-based technology investigated in the second phase was the multi modal presentation of instructions to execute a procedural task. Procedural instructions can be conveyed in a computer-based environment to the learner in several forms, for example, still photograph, video still, line drawing, full motion video, animation (Najjar, 1995a). These media formats can be presented with voice narration or with text caption. This type of presentation in computer-based instruction is normally termed “multimedia”. One of the aspects of the research undertaken in the second phase was the investigation into the presentation of text and sound instructions on a video or animated demonstration. An important aspect of training procedural tasks with computer-based emerging technologies investigated by the second phase of experiments is to ascertain whether the heuristics for learning declarative information from a video clip would be similar, different or the same for learning a procedural task. Preece (1993) makes the point that multimedia has been promoted as an educational tool that can enhance learning and allow users to interact with information in novel ways. However if these sophisticated teaching materials are not designed correctly they may produce memory overload, divided attention or other psychological problems. A specific multimedia configuration may not match the different ways in which people work or learn. The following sections give an

overview of the component parts of multimedia and outline some of the important issues in the development of these elements. These sections also explain why certain of these elements in multimedia have been studied more than others.

5.8.1 Video technology in multimedia

Chapman & Chapman (2000) reveal that video puts a great deal of strain on the processing, storage and the data transmission capacity of a computer. One major consideration when using digital video is the size of the files that are created. The video file comprises of a number of frames, each of these frames is a simple image created by digitalizing the time varying signal produced by the sensors in a video camera. Bitmapped images are used for video frames since this is the format that video equipment produces. The size of the image created for each frame of PAL video, the system used in Western Europe is 768 pixels wide by 576 pixels high at 25 frames per second give 31 Megabytes per second or 1.85 Gigabytes for each minute. Such rates limit the amount of video you can put on a CD-ROM, hard drive or transmit over a network. The high volume of each frame is the result of the storage conditions of bitmapped images. These requirements may be reduced by applying compression to the images, this compression may take several forms, for example reducing the size of the frame or limiting the volume. Ahmad, Akramullah, Liou, & Kafil (2001) indicate that this compression is another production technique that requires a great deal of computing power and high performance machines. These compression methods may result in a significant loss of picture quality. Dick (2002) points out that the constraint with non-linear editing was the technological restrictions of earlier computers. These computers lacked the processing power, data transfer rates

or storage capacity to manipulate large video files. Modern computers have faster processors; larger allocations of memory and large, fast hard disks that have begun to solve editing and storage problems. However there are still advances to be made before all computers will be able to display video that is fast moving with a high resolution.

Video has become popular with both users and designers alike; however the assumptions the user has about the quality of video viewed on a computer will be based on their familiarity with broadcast television. Video is an area of multimedia that is most susceptible to technical change. Vaghan (1998) points out that the video clip is a comparatively recent addition to the multimedia elements and is continually being perfected as the transport; storage, compression and display technologies are being constantly improved. Despite this there are currently several problems with video clips used in multimedia; they are often played back at reduced frame rates and may appear jittery due to dropped frames viewed in small windows on the computer screen and the clips will often exhibit indications of compression (Chapman & Chapman, 2000). Due to these technological restrictions video is often used in moderation in low-end platform multimedia packages.

5.8.2 Sound in Multimedia.

Barron & Atkins (1993) recount that up to the early 1990's the majority of computer-based training programmes concentrated on the visual presentation of information and largely excluded auditory information due to hardware restrictions in computer technology. Historically the introduction of digital audio in 1982 created the conditions for the incorporation of synchronised audio into computer-based

instruction courseware. The advent of audio cards assisted the developers of computer-based instruction to store, edit and play back pieces of audio. An advantage of audio is its flexibility, new sounds or dialogue can be slotted into an existing programme or new files can replace unwanted files. Bradey & Henderson (1995) state that the improvements of computer-generated sound in the nineties redressed to some extent the neglect of the sound channel in multimedia. As noted by Barron & Kysilka, (1993) the increased availability in the 1990s of reasonably priced, good quality digital audio computer cards and computers with sound capabilities permitted computer-based instruction designers to incorporate sound into their programmes.

Beccue & Vila (2001) point out that prior to the appearance of sound computer technology, research had concluded that using multiple senses rather than a single sense had a better outcome for learning. The cost of hardware for the provision of digital sound has been significantly reduced in recent years allowing the wider use of sound in computer-based instruction. Jasper (1992) points out that the fusion of audiovisual instruction and computer-based learning was a later development in multimedia. Before they were brought together they were independent methods of presenting learning materials and were representations of separate traditions.

Audiovisual instruction had its roots in communication and computer-based instruction was derived from computer science. The new advancements in technology has meant that film and video technology has been amalgamated with real and synthetic sounds in a computer environment that was previously restricted to graphics.

Chapman & Chapman (2000) state that there are important differences between sound and the other digital media that constitute the multimedia experience. The

other media are visual, being perceived through the modality of vision whereas sound is sensed through hearing. Sound is a familiar daily phenomenon, however like colour it is a mixture of complicated physical and psychological factors. For example sound may not always be needed to communicate information in a computer presentation and it can also become irritating to the user, especially if noises or sounds are continually repeated. The two types of sound that are most commonly used in multimedia are music and speech. The cultural implications of music and the linguistic content of speech means that these two diverse types of sound work in different ways and are used for different purposes to one another and to other sounds and noises. Dick (2002) makes the point that the uses of sound in computer-based instructions are varied. Musical introductions and background soundtracks are pervasive in multimedia presentations. Sound may also be used to give a warning for time limits or feedback for right or wrong answers in a computer programme. With the appearance of more advanced sound cards, sound can be used for verbal instructions and this is becoming more prevalent in computer-based instruction. The process of listening to sound is very complicated and the computer's sound card has to recreate natural aural experiences as realistically as possible. Sound cards with audio reproduction software have improved a great deal from the early versions.

Beccue & Vila (2001) remark upon the fact that there is a paucity of research on audio in multimedia. There are not enough substantial research findings in this area to develop firm and universal guidelines for applying audio to multimedia in computer-based training. A number of multimedia developers only use sound because new computer technology gives them the opportunity to do so. Others may use it for practical reasons such as attracting the user's attention or for freeing up screen real

estate by using voiceovers instead of lengthy passages of text. In the latter case a common way to include instructions in a complex graphic display is to use vocal narration (Barron & Atkins, 1993). The results from the few studies that have investigated the inclusion of sound into multimedia are conflicting. Bradey & Henderson (1995) reported a positive attitude of users to voiceovers in computer-based instruction. Barron & Kysilka (1993) found no significant difference between sound and text conditions for learning declarative knowledge on a CD-ROM whilst Mayer & Anderson (1991) recount findings that demonstrated an advantage for an instructional animation with voice narration over an animation with text or an animation with neither. Commentators in the literature, for example Beccue & Vila, (2001) postulate that learner variables such as learning styles and level of computer literacy play a role in determining the efficiency of computer-based instruction with sound. The phase of experiment investigated the placing of vocal instructions on a video clip for learning and retaining a procedural task.

5.8.3 Text in Multimedia.

During the lifespan of multimedia in computer-based instruction there has been more research carried out on text than sound since as previously mentioned the efficient production of audio is a recent addition to multimedia. As far as text is concerned a sizable portion of mainstream multimedia research is concerned with “hypertext”. Chapman & Chapman (2000) define hypertext as text with augmented links that are indicators to other sections of text located in the same or a different document. This form of navigation through pages is associated more with web pages on the internet but is also employed to navigate pages in multimedia presentations.

Another area of that research that affects multimedia is the use of text for the learning and retention of information. A body of research work exists that describes investigations into the use of text with pictures for learning, for example whether pictures support learning from text (Carney & Levin, 2002). Another related research area is whether pictorial information is better remembered when supported with redundant information in the form of text captions for learning scientific explanations (Mayer, 1989). Researchers in computer-based instruction and multimedia have tapped into this large body of theory and research into the use of text with graphics or pictures and applied it to multimedia research.

One argument in favour of the use of dialogue as opposed to text in computer-based instruction has its foundations in the multi-channel theory for information processing as espoused by Paivio, (1986). The representations of teaching materials are greatly enhanced when the learning materials are presented in a redundant fashion through the two modalities of sight and hearing. Another argument for using sound in computer based instruction is that text is harder to read from a computer screen than from paper. According to some studies users take up to thirty percent longer to read text on a computer screen (Gould & Grischkowsky, 1984). Wade & Swanson (2001) point out that as well as being slower, users are also less accurate for activities such as proof reading using a computer when compared to reading text on paper. The manifestation of tired and sore eyes is also attributed to reading text from a computer screen.

However an argument in favour of text on a computer screen is based around the idea that text is better retained over time compared to speech, this is founded on the

concept that a greater cognitive effort is expended in encoding text than other media, increased demand augments processing thus leads to the textual information being better remembered over (Walker, Jones, & Mar, 1983). The second phase of experiments investigated the dimension of learning procedural tasks highlighted by previous studies concerning the relationship between video or television and text or voice narration. Salmon (1984) draws attention to a possible disadvantage of learning with television or full motion video. This researcher tested children's ability to make inferences about information presented in two formats. First as a silent television film, second as a text booklet with comparable information to the film. Results showed a significant number were able to make more inferences in the text condition. The children in the study were more familiar with television medium; however they found using the text to perform the task more demanding. Reading text is generally less popular than visual demonstrations for adults as well (Carroll, Mack, Lewis, Grischkowsky, & Robertson, 1985). Findings from the latter study indicated that learners would rather see a visual demonstration showing how a computer package works rather than read a text manual. Text instructions are processed more thoroughly than viewing a video demonstration of a task. Readers of text instructions must fit a verbal label to an equivalent action or object in the real world. This is known as the referential step (Just & Carpenter, 1987). The user must remember what each component or part of the procedure is and where it goes in relation to other parts. From textual information alone the user must imagine the series of actions. The argument is that the viewer of a visual demonstration of a procedure does not have to make this referential step.

Other research has shown that listening to text being spoken and reading text are processed differently in the human brain. In the area of story comprehension, research has suggested that readers recall more of the surface structure of written text and retain the information for a longer period of time than listeners. Another argument is that reading text is less prone to distortion over time compared to listening to dialogue (Horowitz, 1968). The second phase of the experiments investigated the findings of these above experiments and tried to determine if they apply to learning an assembly task with text in the multimedia materials. There are other issues in multimedia that have a bearing on learning, for example where the text should be placed on the screen and whether it should be placed simultaneously with the graphic material. There is also research that has explored the use of text and animation, for example Mayer & Moreno (1998) compared animation plus text versus animation plus sound for learning declarative information and found an advantage for the latter presentation format for learning.

5.9 Chapter summary.

- A major reason why any guidelines for the construction of multimedia based on declarative learning research may be problematic for learning procedural tasks is that these two types of learning may entail separate memory systems. There exists a body of research that supports the notion of short term and long term memory stores. Further research has produced evidence for a dichotomy of long-term memory. This dichotomy consists of a memory system for procedural tasks and a memory system for declarative knowledge.

- Current mainstream guidelines for the configuration of media in dynamic and static visual displays have been developed by Richard Mayer and his colleagues at the University of Santa Barbara in California. These guidelines are based on three theoretical perspectives; the dual-coding theory, cognitive load theory and constructivist learning theory. The research these guidelines are based on involved studies into configuring multimedia for learning declarative knowledge. Given the evidence that separate procedural and declarative memory systems may exist, such guidelines or heuristics may not be useful for learning procedural tasks.
- There seven guidelines or principles for the design of multimedia learning packages. Some of these principles are more directly at the configuration of media in dynamic visual display than others. The principles tested in the second phase of experiments are; the multimedia principle, the temporal contiguity principle, the, modality principle and the redundancy principle on multimedia. These principles or guidelines are used to create multimedia instructions for a procedural task.
- The three elements of multimedia that the second phase manipulated in multimedia demonstrations in the second phase were; video clips, sound and text. Advances in technology have meant that both video streaming and sound are being used more in multimedia presentations. The memory and hard drive storage in a computer required for video streaming is large. The increase in memory and storage space and the introduction of easier editing facilities has made video streaming available to a growing number of users.

- Sound is also being used more in multimedia design in computing technology.

Voice narrations have become a popular method of giving instructions or information in multimedia learning packages. Text is the most researched element of multimedia due to the fact that sound and vocal narrations were less common in the early days of computing. There is evidence from studies that text on a computer screen is harder to read, however information that is read is retained longer.

6 Design and methodology of second phase of experiments.

“The best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition.”

Clark (1983, page 445)

6.1 Focus of second phase of experiments in research programme.

The decision to focus the next phase of experiments on learning and retaining procedural tasks required a different methodology from the first phase of experiments.

The second phase concentrated on the best way to train a learner to perform a procedural task in different multimedia conditions. Whereas wearable computer and display technology may be in constant flux and development, the multi-modal training materials will remain a necessary requirement for computer-based training.

Knowledge concerning the optimal method of presenting media in multimedia demonstrations of procedural tasks is fundamental to any hardware vehicle used for training. The experiments in phase one of the research programme were tailored towards identifying differences between hardware types for following procedural instructions. The experimental paradigm used in these first experiments did not assess the retention of the assembly tasks in different multimedia conditions.

The second phase of the experiments focused on the training of procedural tasks on a video clip of the type that would be used in any computer based technology. The use

of the video clip has become ubiquitous in computer-based instruction, not only in desktop computers but also in emerging technologies such as wearable computers and hand held computing devices. Research into using video clips for training procedural tasks is limited and not as advanced as research into using animations for teaching declarative information. Research into learning declarative facts using multimedia has created a body of knowledge relating to the optimum configuration of multimedia information. The main research question investigated in the second phase of the experiments was whether the heuristics and guidelines for using multimedia computer-based instruction for declarative knowledge could be generalised to learning procedural tasks. The main difference in methodology between the first and second phase of experiments was that the experiments in the second phase comprised of three experimental sessions, a training session a criterion session and a retention session to test the effectiveness of several multimedia configurations.

The procedural task in the second phase was similar to the first phase. The task again involved the assembly of Lego models, however in these experiments the models were recognizable objects rather than abstract models as used in the first phase. This type of procedural task was used since assembly and disassembly are subtasks in many procedural tasks, for example changing oil, replacing brake pads and installing equipment (Ellis, Whitehall, & Irick, 1996). Actual artefacts, models of a helicopter, plane, car and pick up truck were used in the experiments so that participants could form structural relationships between the parts and have a better chance of building the models from memory.

This chapter describes the experimental paradigm used in the second phase of experiments. The chapter then goes on to outline previous research done with assembling model kits and with video demonstrations for procedural tasks. There is an overview of previous studies and the possible methodological differences between them. The next section outlines the theoretical basis for the four experiments in the second phase of experiments. The possible existence of a taxonomy of procedural tasks is then discussed, this may explain conflicting findings in studies of the configuration of multimedia materials for learning procedural tasks. Lastly this chapter looks at the hardware used in the second phase of experiments. After continued problems with the head-mounted display in the first experiment the second phase, it was decided to use a desktop computer for the last three experiments.

6.2 New experimental paradigm.

The first phase of experiments tested the ability of each participant to follow instructions in three conditions and build the model using 3 different methods of instruction. The second phase of experiments focused on learning to build a Lego model in different multimedia conditions using one instruction format. Salmoni, Schmidt, & Walter (1984) give the most commonly accepted definition of learning as “a relatively permanent change resulting from practice or experience in the capability of responding.” Palmiter & Elkerton (1993) make the point that when the training or learning method is to be evaluated, the assessment of information learned must exceed a basic acquisition phase. Schmidt & Bjork (1992) state that acquisition performance is not an appropriate guide to learning. Evaluation of learning at some future date is necessary to evaluate the initial training session. When studies of learning take place researchers normally ask participants to practise a task in an

acquisition phase where a particular independent variable is manipulated. This independent variable can be the type of instructions, feedback or scheduling of practise. The performance on the task is normally recorded using practise trials of the participant groups learning with different levels of the independent variable. The reason for this type of experimental paradigm is that the conditions that increase the rate of improvement or cause participants to reach criterion more quickly are assumed to be the most efficient for learning the task. There are two related problems with using this system of evaluation for learning.

The first problem is that measuring the acquisition performance itself is not the optimum gauge for learning. In recent decades research on the methods of learning, memory and performance has neglected the important difference between the short-lived strength or accessibility of a response and the practiced or habit strength of that response. Experimental variables applied during learning can have two distinctive types of effects. The first effect is that these variables can have an enduring change in behaviour; they may quicken the development of a specific ability so that participants with an increase in this capability will outperform the control group. The second effect manifests itself in temporary changes in behaviour as a result of the manipulations that take place in the experiment. These effects may improve or diminish performance and may disappear or alter significantly when the manipulation is removed. These performance effects can be caused by many factors. Performance may be momentarily enhanced by motivating instructions or by attempting to please the experimenter. On the other hand performance may be poor due to boredom or fatigue. This basic distinction was common knowledge in the behaviourist paradigm from the middle of the last century for example (Estes, 1955; Hull, 1943).

Schmidt & Bjork (1992) speculate that the information-processing model typical of cognitive psychology, which has dominated research at the end of the last century, is responsible for theorists gravitating away from this important distinction between the two learning effects. The information-processing model is based on the architecture of a digital computer and this model does not assume the existence of these two types of memory performance. The difficulty, however is to determine which of the practice variables has produced learning; more importantly has this independent variable created permanent or temporary learning. In a majority of learning scenarios the aim of the practice session is to furnish the learner with the ability to remember the learning sometime in the future (Salmoni et al., 1984). The way around this problem is to use retention tests as a method of testing the extent to which proper learning has occurred. A retention phase is added after an interpolated space of time that is long enough to ensure that the transitory effects of the independent variable have disappeared. When participants are tested in the same task with the same independent variable some time later, performance differences in the same groups can be regarded with more certainty as indications of different learning compared to that which took place in the acquisition phase.

The second problem with evaluating learning is that the acquisition and retention phases of the experiment involve different types of learning. Learning in the training session is considered to be the processes which occur during practice whereas learning in a retention phase involve a different set of mental processes that take place some time after practice. Since acquisition and retention of learning are considered to be distinct phenomena, researchers often treat them like different methods. Despite

this the level of retention should be seen as replication of the acquisition phase and should therefore be studied in the same conditions as the acquisition phase of the experiment. Schmidt & Bjork (1992) view the retention phase as the true indicator of learning. The amount learned in acquisition is not seen as a reliable measure therefore the retention phase should take precedence in measuring the learning.

The experimental paradigm used in the second phase of experiments was adapted from a series of experiments by (Ellis et al., 1996). Their study involved assembling a model crane with picture and text instructions. The experimental procedure consisted of three separate assemblies, the initial assembly, the criterion assembly and the retention assembly. The experiments in the second phase of experiments largely followed the above paradigm but substituted the text and still pictures with videos of hands building a model in different multimedia conditions. Participants watched the video clips on Windows media player. In the initial assembly participants were shown how the models were built and they then built the model themselves viewing the instructional video if they needed to. In the criterion phase the participants attempted to build the model without accessing the video clip at all to make sure they could build the model to criterion. In the retention phase the participants returned a week to ten days later and built the model again but this time they could use the video as a guide if they were stuck or unsure of the assembly.

The advantage of the multimedia presentation may be associated with the amount of time elapsed between the viewing the multimedia and the testing of knowledge gained from the presentation. There may be different results for testing several weeks after the presentations compared to testing after several minutes. Peek (1989) states that

testing participants directly after viewing multimedia presentations is a methodologically unsound practice. Immediate testing does not provide the experimenter with data about the longer-term retention of the information in the multimedia presentation. A more efficient method would be to test immediately or in a training phase then perform the same test after a delay of at least several days to assess retention in multimedia studies. (Levie & Lentz, 1982) illustrated this point in a literature review, this review demonstrated that in nineteen comparison studies, pictures facilitated the memory for text more in delayed recall than for immediate recall. The average advantage due to pictures was reported in the above review as five times larger in delayed testing than for immediate testing.

6.3 Previous Research on Assembly tasks using Model Kits.

In the 1980s several researchers investigated aspects of learning an assembly task using commercially available model kits similar to Lego. Stone & Glock (1981) carried out a seminal study that investigated the way in which people read and use directions or procedural information. The experimental task in their study involved participants assembling a Fischer model car. The directions given to the participants were either text instructions for the assembly on their own, illustrations of each stage of the construction and finally with both the text instructions plus the illustrations of each stage. The results of this part of the experiment revealed that the text with illustrations produced significantly more accurate mean performance than the other conditions. This study also analysed the eye movements of the participants as they used the instructional materials. Video cameras simultaneously recorded the subjects' hands and the subjects' head to ascertain the way in which they built the model and when they looked at the illustrations or the text instructions. A projector displayed the

illustrations and the text instructions were presented separately on a paper. The results of this study showed a significant advantage for the text plus illustrations condition. The post-test analysis of the video filming of the participants showed that in the text plus illustration condition participants spent four times as long looking at the text than the illustrations. This study was important for the design of multimedia procedural instructions since the findings supported the idea that visual information plus text enhances performance of following directions. However the filmed behaviour of the participants in the illustrations plus text condition indicated that the text instructions were referred to more than the illustrations.

A series of studies carried out by Patricia Baggett in the eighties investigated the experimental designs to augment learning and retention of assembly tasks using multimedia training materials (Baggett & Ehrenfeucht, 1982 ; Baggett, 1986, 1987). In one study Baggett (1987) used a film with vocal instructions to teach participants to assemble a Fisher Technik model kit of a helicopter. This experiment set out primarily to test the optimum order in which to present the film with practice schedules. What was unique about this experiment was that the researcher placed a great deal of importance on giving the kit parts particular names in the instructions based on their appearance. These names were purported to help the participants remember the assembly of the parts. The names that were selected were short and simple and supposed to correspond with the physical qualities of the parts. The naming procedure was done scientifically using a technique described by (Baggett & Ehrenfeucht, 1982).

Ellis et al. (1996) explored the training and retention of an assembly task using a model kit. This experiment had an initial training phase of the assembly task followed by an assembly test. Participants came back a week later and were tested again on the assembly task. The experimental task was to build a forty-six-piece Capsella motorised model kit of a crane. Participants were trained to build the model in four conditions: with structural text instructions (where each part goes in relation to another), structural text instructions plus a picture of the completed model, functional text instructions (what the function of each part is in relation to another part) and functional text instructions plus a picture of the completed model. The main objective of this experiment was to discover if the functional or structural instructions had an advantage for the retention and transfer of the assembly task. Results showed an advantage for the retention but not the transfer of a task with functional instructions. Structural instructions aided both retention and transfer of skill to another task. As far as the picture plus text versus the text alone element of the experiment was concerned, results showed that the picture condition had an advantage only until the assembly was learned to criterion and the picture had no affect on performance after that.

6.3.1 Rationale behind research into video demonstrations.

The second phase of experiments focused on aspects of the training materials used in computer-based instruction. Park & Hopkins (1993) relate that the ability of computer technology to manage multimedia information has meant that some form of visual dynamic presentation has become a primary component of media-based instruction. The visual presentations in multimedia instruction are normally animations of artificially created still pictures or full motion video of real sequences or events. The dynamic aspect of animation and video bring three features to a learning environment:

visualisation, motion and trajectory of an object. Visual material with motion is valuable for representing actions and ideas that change over time. Representing tasks involving motion with animation or video activates the learner's automatic ability of their visual system to induce apparent movement and encode this motion directly into memory. Static video stills or pictures demand more effort to create a mental model of the dynamic nature of the task through incorporating and linking individual items of information.

Spangenberg (1973) provided evidence that a video film to was more effective than video stills. In this study mechanics that were taught how to disassemble a M85 machine gun using the video were quicker and made less sequence errors than mechanics in the video stills condition. The explanation for this finding was that the motion in the video attracted attention and functions as a cue to the critical elements of the procedure or direction of movement. Dwyer (1982) notes that the main rationalization for using visual materials in the learning process is to enhance or simplify points and insure the accuracy and regulation of the information being conveyed. Visual materials have more learning cues so learning is augmented by visual presentations. It is believed that for the purpose of learning visual illustrations are not multi-purpose in nature, some may be more effective than others in aiding learning for particular educational objectives. Proponents of the single channel hypothesis of information processing, for example (Pylyshyn, 1981) would argue that the limited capacity of the information processing of the perceptual systems precludes the addition of cues or very realistic cues; these may be distracting to the learner. There may be a learning situation where processing accurate detail past a particular point contributes little or might even impede learning. On the other hand advocates of

multi channel information processing, for example Mayer & Moreno (1998) argue that research into learning with multimodal animations has facilitated learning declarative information or aided transfer of knowledge for problem solving.

In a review of the literature on instructional materials with moving images Park & Hopkins (1993) make the point that much of the early research was carried out on films and television. However, due to the advance in computer technology and the ability of computers to handle animation and video, much of the recent research has concentrated on computer-based instruction. Large (1996) relates that most of the research on moving images in computer-based instruction has dealt with animation, since this medium is more common than full motion video. The reason for the overrepresentation of animation studies in the multimedia literature is that until recently video had a higher demand on the storage and retrieval capabilities of most computers. Text remains predominant for transmitting information in the instructions. This has been due to the slower development of sound in computers (see section 5.8.2).

Alty (1993) is of the opinion that one of the most important factors in designing a multimedia interface is the type of information users need to enable them to carry out tasks. A major aspect of the design process is deciding which medium best conveys the information. Park & Hopkins (1993) indicate that previous research has shown that training materials with motion have an advantage for a procedural task, for example operating or repairing equipment or assembling artefacts, because the sequence of steps or procedures can be explicitly demonstrated. A static presentation cannot illustrate the actual motion involved in the procedure. Using a film or a video

of a procedural task is seen as a cost effective method of replacing a live demonstration from an expert. These arguments for using full motion video as a medium for learning procedural tasks coupled with the recent improvements in computer memory may well mean that video will be increasingly used to demonstrate procedures in computer-based instruction. The arguments in favour of using video as opposed to animation are based on the dynamic nature of video.

The wearable computer with a head-mounted display is not the only emerging technological device that will use full motion video for training purposes. Sharples (2000) describes the concept of “lifelong learning” and its use of developing technology. Life long learning is an approach to learning based on the need to provide the population with the skills and knowledge to be successful in a constantly changing work environment. One development that will enhance life long learning is that in the near future learners will not be tied to specific locations. Emerging mobile computer technologies will allow them to study at home, at the workplace, in a library as well as further education establishments. They can study in different locations and at their own convenience using portable mobile devices. People will also have the propensity to study from a distance using broadcast media and the internet.

As learning has evolved into an individual learner-centred process the new computer-based technologies have developed into more personal devices. Computer technology has become ubiquitous and appears in many modern appliances that seek to substitute or support human action. Whereas the computer hardware may be ephemeral and become outdated the software packages evolve through consecutive versions with a large amount of backward compatibility (Barber & Baumann, 2002). This allows

learners to maintain and build on previous knowledge gained from the training software. A consequence of the new personal technology is that it offers the opportunity to create and design personal specialised learning materials. The hand held computer or PDA could feasibly rival the head-mounted display as a vehicle for learning declarative and procedural knowledge.

An important consideration in the development of mobile learning devices is whether they should be a software system with the ability to run a number of devices or a fixed amalgamation of software and hardware. It may be that in order to last over a period of time the mobile learning system should be a software environment with the ability to run on diverse hardware platforms such as wearable computers, desktop computers or palmtop devices. This would circumvent obsolescence caused by technological advance as well as allowing the learner to use whatever hardware they find most convenient.

An example of the ongoing advancement of mobile computer devices is the tablet personal computer. This device is a cross between a laptop and a P.D.A. (Krakow, 2002). This mobile device is similar to a laptop in size and allows digital pen and ink input like a P.D.A. The tablet personal computer is also capable of voice recognition and some versions will incorporate a wireless ethernet. Lee & Owens (2000) predict that a majority of computer-based training software used in current and emerging hardware will have a strong emphasis on multimedia with video used as a demonstration medium.

6.3.2 Previous Research with Video Demonstrations.

In the past twenty years there have only been four studies, which have examined retention and performance of procedural tasks from memory using an dynamic visual demonstration. Three of these studies employed actual filmed demonstrations of procedural tasks. Two of these studies examined adults learning the HyperCard authoring system and the third used a video clip demonstrating the correct method of bandaging an arm. The fourth study tested the effectiveness of a sixteen colour animation for showing twelve year old children how to find south using the sun's shadow. This section will concentrate on the three studies that tested the effectiveness of filmed demonstrations in comparison to other instruction formats.

In the early nineties there were several studies that compared filmed demonstrations to text instructions. This line of research was initiated by evidence from the nineteen eighties and nineties that instruction manuals for computer-based software and hardware were unpopular with users. Carroll et al. (1985) observed that users prefer to learn by trying out a system themselves rather than reading through a complex instruction manual. Mack, Lewis, & Carroll (1983) are of the opinion that training problems arise in the office and in work environments because the written instructions to learn a computer application or new piece of technology are demanding to work through and take time to assimilate. The average learner would prefer to have an expert or a colleague to demonstrate how to use the technology or perform a task. However tutorial help from trained experts can prove to be costly and time consuming. These observations of training difficulties led to a plethora of “watch

me do it” demonstrations being incorporated into learning interfaces or computer packages in the mid to late nineteen eighties.

Palmiter, And Elkerton, & Baggett (1991) carried out an investigation into the effectiveness of a filmed demonstration compared to text instructions for teaching users how to use a computer database. The above authors considered some of the perceived advantages and disadvantages of filmed or animated demonstrations. For example animated demonstrations may convey procedural knowledge about the task more precisely. The animation demonstrates to the user how the task is executed as the steps of the procedure are carried out. Learners should be able to go through visually how each procedural step plays a role in the whole task. Animated demonstrations may permit learners to rehearse and plan in visual terms whilst watching the instructions being demonstrated. One supposed advantage for animated instructions is that it increases preliminary learning in comparison to text instructions due to a reduction in the level of cognitive processing during the learning stage. As far as learning to use an interface or a computer package is concerned, a reason for the reduced processing for animated demonstrations in the learning phase is that the demonstrations seem more like examples than written procedures in a manual. The demonstrations appear to be representations of the task that has to be learned and therefore serve as example and could in fact be compared to written examples.

Previous research has shown that well constructed examples improve learning in different knowledge domains. When they have the choice between using examples or written instructions, learners tend to rely on examples and process written instructions superficially. Lewis & Anderson (1985) point out that the users’ reliance on examples

and performance improvement appears to have a close correspondence between the features of the problem and the example given to solve the problem. This may be the case for filmed or animated demonstrations for learning procedural tasks. The demonstrations act as visual examples of ways in which the interface or software can be used. The contention here is that learners will be more likely to use animated demonstrations and find this form easier to use. Palmiter et al. (1991) foresaw a possible disadvantage of animated demonstrations. That is that the learners may passively watch the demonstration of the task with a minimum of processing and encoding of the procedure that had to be learned.

Palmiter & Elkerton (1991) compared filmed visual instructions to the same material in text form for teaching the HyperCard computer application. This visual demonstration was without a narration and showed the procedures for using the interface and the sequence of actions with the menus and dialogue boxes for executing specific tasks. The text instructions were based on the help systems bundled with the computer package. These text instructions were presented to the learner on the computer screen. The study had an initial training period followed an immediate test session. The participants then returned three to seven days later and performed the same tasks as they were taught in the training session. The results of this study indicated that the demonstration group outperformed text instruction group in the initial training session for speed. However in immediate test and retention session of the experiment the mean times for the text instruction group improved considerably and were on average quicker than the demonstration group. When the percentage of correct trials was calculated for both conditions the text group had lower percentage of trials correct in the training session but had roughly the same number of trials

correct at immediate test as the demonstration group but improved further in the retention session with a higher percentage of trials correct than the demonstration group.

These findings were taken as a vindication of the idea that filmed demonstrations are not processed as well as text instructions. The demonstration reduces the amount of translation necessary when training starts and acts as an interface example.

Participants made fewer mistakes initially in the demonstration condition because they simply mimicked the film without fully learning the procedures. The text group were slower and made more mistakes in the training phase since reading the instructions took more cognitive effort. The text instruction group was exposed to a more detailed encoding medium since it contained visual, verbal and motor information, whereas the demonstration group were only presented with visual and motor sequences on the film.

In a follow up study by (Palmiter & Elkerton, 1993) a verbal component was added to the filmed demonstration of the same task to test whether this would improve the performance of a demonstration group in tests after task acquisition in the initial training period. It was hypothesized that by adding a verbal component this may cause the participants to process the instructions at a deeper level rather than passively watch the demonstration then mimic it during task training. This study replicated the one carried out by (Palmiter & Elkerton, 1991) except that more tasks were included and a third condition was added, this comprised of a combination of the text only and demonstration only conditions. In this third condition matching narrated instructions were incorporated into the demonstration. A voice narration was added rather than on

screen text due to the practical difficulties of placing text simultaneously with the corresponding visual instructions. The researchers thought that placing the text instructions on the demonstration would lead to a busier display that may confuse the viewer. Moreover the text on screen would increase the amount of time needed to watch the demonstration.

The researchers used the concept promoted by (Baggett & Ehrenfeucht, 1983) which states that by combining the demonstration with narration could reduce competition for attending to two sources of information simultaneously. The results of this study indicated that the two demonstration groups were significantly faster and more accurate than the text instructions only group in the immediate test after the training phase. However the two demonstration groups consistently became slower between the post training test and the delayed testing session a week later. On the other hand text instruction users performed at a steady pace between the immediate test and the delayed testing sessions.

An unexpected result from this study was that the demonstration plus voice and the demonstration only groups had very similar results throughout the study. The authors of this study conclude that adding voice instructions is not an adequate solution to the poorer retention of a procedural task learned from a filmed demonstration. The authors give further reasons for the fast and accurate performance of the demonstration groups in the initial phase compared to the text only group. The demonstration groups did not require to map the verbal ideas to items and actions in the interface or determine the position of items in the interface. The demonstration groups were quicker in the training phase because they did not need to spend time

figuring out the steps the text only group read. Participants in the demonstration groups were able to see the way the system responded to user input so they were not surprised or distracted by the consequences of their actions. These factors were seen to have led to a more homogeneous performance of the two demonstration groups in the initial training phase. Questionnaire feedback from participants showed that learners in the demonstration groups felt that they simply mimicked the demonstrations and this superficial processing would explain why they took less time to acquire the procedures. The authors conclude by noting that despite the fact that they were less effective for the retention of procedural tasks feedback from questionnaires indicated that filmed demonstrations were more popular with users.

In a third study Michas & Berry (2000) examined the efficiency of the filmed demonstration compared to other types of instructions. The procedural task participants had to learn in this study was the sequence of seven steps for the correct bandaging of an injured arm. The study comprised of three experiments that compared several instruction formats; text alone, line drawings, line drawings with captions, video stills and full motion video. The study set out to test whether some of the multimedia heuristics and guidelines created for learning declarative information were applicable to learning a procedural task. They tested some of the theories espoused by Mayer and colleagues for learning and transferring declarative knowledge. The first experiment in this study looked at combining a visual and a verbal component in the instructions and had five conditions; text only, line drawing only, video (with only a visual component), line drawing plus text and video stills. This experiment set out to study the concepts about multimedia presentations produced by (Mayer, 1997) and the (Paivio, 1986) dual coding theory of learning.

The experiment included an initial learning phase in which learners had to study the learning material until they felt confident enough to perform the task of bandaging an arm correctly without the benefit of the instructions. Participants when ready performed the bandaging themselves, were judged on this performance then answered questions about the process. The results of the study showed that the text plus line drawings instructions and the video instructions conditions were not significantly superior to one another but were significantly better than the other three formats for learning the bandaging task. The authors concluded that this result indicated that learning a procedural task does not require both the verbal and visual channels for processing the instructions since the video film without text or voice instructions was as effective as the line drawings with text. These findings contradict the dual coding theory and the work of Mayer and Paivio.

A second experiment in the (Michas & Berry, 2000) study demonstrated that enhanced drawings with arrows providing action information accompanied by text instructions produced better learning results than enhanced drawings on their own, line drawings with text instructions or line drawings on their own. The last experiment in this study tested the “temporal contiguity principle (Mayer, 1999), which states that there is an advantage when text and matching visual instructions are placed together on the same screen. The results of the above experiment showed no significant difference in remembering the task between conditions where text was placed simultaneously, placed apart or separated in time from illustrations. The authors concluded that text instructions and corresponding illustrations for learning a procedural task need not be presented contiguously in space and time. These results

constituted a seminal challenge to the use of guidelines for constructing multimedia instructions for instructions for procedural tasks based on the dual coding theory and the cognitive load theory.

6.3.3 Methodological differences between studies.

As Large (1996) recounts much attention in this media research domain has focused on whether animation or moving visual displays improve the recall of textual information and on the optimum method of designing multimedia materials to augment such improvements. Results from studies involving computer-based instruction packages have been contradictory and equivocal. Some studies into the use of animation accompanied by text for example, demonstrated an advantage for animation plus verbal information whilst others did not. Park & Hopkins (1993) generalised such contradictory findings to studies that compared all dynamic visual displays (animations, video and television) as well as still photographs and line drawings. Large (1996) points out that a major problem with research into learning with multimedia is the number of possible variables involved. The configuration of the multimedia presentations can vary according to a wide range of criteria in studies. Some of these criteria related to dynamic visual displays are affected by hardware as well as design considerations. The lucidity, size and quality of the image will vary from study to study. The duration of the video or animation and the amount of learner control are another two decisive factors that will differ. The verbal information that accompanies the visual display is also prone to variation in studies. Text for example can fluctuate in length, level of complexity and layout. Text and voiceover can be presented as narrative, descriptive, procedural or an amalgam of these categories.

Another source of variability is the type of user of multimedia, for example adult learners and school children may react differently to multimedia instruction. Other well-documented user variables include intellectual ability, spatial skills and prior knowledge of the subject domain.

Variability can also occur between studies for the method in which recall of the multimedia is being tested. Some researchers, for example Large, Besheti, Breleux, & Renaud (1995) asked participants to recall in their own words what they have learned, this feedback was then re-analysed to measure the participants ability to draw inferences from the information. In other studies participants have been requested to fill in multiple-choice questionnaires or complete a problem-solving task to gauge understanding (Large et al., 1994). Large (1996) argues that spoken or written tests may favour the recall of verbal elements of the multimedia rather than the information that is conveyed visually.

6.4 Taxonomy of Procedural tasks.

The second phase of experiments investigated the learning and retention of procedural tasks. Konoske & Ellis (1991) define procedural tasks as ordered sequence of steps or actions performed on an object or in a particular situation to achieve an objective. Procedural tasks involve few decisions points and are usually executed in the same fashion each time the task is performed. Procedural tasks are seen to vary in the required planning to accomplish the task, the amount of cueing incorporated into the task and the number of decision points needed. Other variables in procedural tasks include; the number of steps that are performed and whether the aims of the procedural task are internal or external to the task. The goal of an operator task, for

example is external to the task, situation or system, however the objectives of a maintenance task would be internal to the task or situation. Employees in technical occupations tend mostly to carry out procedural tasks in their day-to-day work. However research has indicated that these procedural tasks are not very often well retained (Hagman & Rose, 1983). The research undertaken in the second phase sought to discover the particular configuration of multimedia presentation of training materials that would best improve retention of these procedural skills.

Konoske & Ellis (1991) envisaged a taxonomy of procedural tasks that reflect differences in the characteristics of several sub-types of procedural tasks. The first type of task is the “operator task”, this requires the performance of a procedure to acquire an objective external to the system or the device. The system or device is used as a tool or a vehicle to accomplish that goal and the means by which that goal is achieved. An example of an operator task would be driving a car or operating a machine in a factory. The second type of task in this taxonomy comes under the rubric of “maintenance/repair/assembly task. This type of task requires the repair, maintenance, assembly or disassembly of a device or a system. Maintenance and repair tasks involve a particular action on the device or the system. Examples of these tasks would include fixing a bicycle brake or changing the oil in a car. Assembly tasks require constructing an object out of component parts. The third type of procedural task in this taxonomy is the “paper based task” which involves the use of a particular arrangement of procedures for preparing documents or completing forms in the correct format. The fourth type of procedural task is “locating objects or information task” This task involves locating information or objects that are available in a database or similar type of repository. An example of this procedural task would be using a dictionary or retrieving a part from a store. A taxonomy of procedural tasks

may explain some of the contradictory results in previous studies involving learning procedural tasks with multimedia materials. Each subtype of procedural task may require a different configuration or presentation of media.

6.4.1 Declarative elements in procedural learning.

The study by (Baggett, 1987) mentioned earlier in this chapter (see section 6), investigated learning to build a model kit from multimedia instructions. This was part of a series of studies undertaken by Baggett and colleagues that explored the use of media with both declarative and procedural learning by using films with narration as the media. However in their research these investigators do not make the distinction between declarative and procedural learning but treat all learning as homogeneous. For example, Baggett & Ehrenfeucht (1983) experimented with the presentation of narration sequentially and simultaneously on an educational film about Venus flytraps. In another study Baggett (1984) tested the dual-coding theory using a film that introduced participants to an assembly kit, its pieces, and names of the pieces and the uses of the uses of the pieces. In this study there were several conditions where the narration gave the information verbally at different time intervals before and after the visual information was presented as well as simultaneously with the visual elements in the film. A major research concern of these above studies was testing the dual coding theory for the learning and retention of declarative visual and verbal materials on educational films (Baggett, 1989).

Baggett & Ehrenfeucht (1982) developed an empirical method for the creation of names for unfamiliar pieces or parts of a kit. These parts, unlike a “wheel” or

“windscreen” are not recognizable as everyday objects with familiar names that would aid memory for the piece. In order to generate suitable names that best describe these objects, a standard procedure involving three principles was devised. Three principles or guidelines employed in deriving the names were: the vocabulary and the construction of the names should be within the parameters of the users linguistic abilities; the names should be economical in terms of efficiency, that is short but distinctive; and lastly that names should create a classification system. This means that a name must include a generic term and when needed one or two modifiers. Normally in this naming system the generic term is a noun and the modifiers are adjectives and prepositional phrases.

The technique for generating names according to these principles entails a three-stage procedure where first, a group of participants generates names; second, from the names generated the experimenter selects a subset of the names that correspond to the following criteria, if a particular name is chosen the most often, the name is the shortest and the names stay within in the classification system set by the naming procedure. The names are tested in two ways, first by evaluating how well participants matched the names with the corresponding piece and second by evaluating how well the participants could recall the names when shown the actual objects. The two latter stages of this process were iterated, that is if a particular name is poorly matched or remembered then it is substituted by another generated name then tested in the same manner. This method according to its creators produces names that develop a classification system and the names are natural, short and associated with their physical characteristic that were easy to recall. This naming system was used in the Baggett (1987) investigation into the effects of practice on building a

model kit. This study investigated learning to build a model helicopter in conditions that involved different levels of practice. This study did not manipulate the media in the film in the different conditions, nor did the participants interact with the film since this study entailed watching a film on a distant screen so there was no opportunity for interacting with the learning materials or evaluate constructivism since this was not a research concern. What makes this study interesting was that the names given to many of the pieces used in the film narration were specially created to enhance the recall of these particular pieces.

The presentation of the media in the film used the (Paivio, 1986) dual-coding theory as a guide for presenting the voice narration and did not consider the possible difference between declarative and procedural learning. Using the sophisticated naming procedure as described above indicates a strong belief by the researcher that this type of declarative information would contribute to learning how to build the model helicopter. There is no apparent theoretical perspective that posits that memory for assembling an object will be enhanced by making the names of the component parts of the model easy to remember. Baggett & Ehrenfeucht (1983) clearly presupposed that the media instructions for learning factual and procedural information should be configured in the same way. The second phase of experiments in this thesis used a naming procedure adapted from and based on the above scientific procedure for naming the model parts that did not have clearly identifiable names. Participants were tested in recall of these names in order to investigate whether this was information that learners processed in any depth whilst viewing the video clip showing the model being built.

6.5 Hardware used in second phase of experiments.

The general features of the optical systems in head-mounted displays may present problems for viewing information and these characteristics may confound experiments in multimedia. The first experiment in the second phase used the Sony Glasstron head-mounted display in opaque mode for viewing the video demonstrations. The Glasstron was used in the second phase because it outperformed the Albatech Personal Monitor since the Glasstron uses a SVGA computer signal. A head-mounted display was used in the first experiment in the second phase in order to give a sense of continuity in the research and build on earlier qualitative results from the first phase where a majority of participants found it useful to switch their gaze from the screen to the assembly task using the head-mounted display. However feedback from the post-test interviews after the first experiment in phase two indicated that some participants reported visual problems using the head-mounted display and reported that they could not see the screen clearly. This feedback resulted in the reconsideration of the use of a head-mounted display in the remainder of the experiments.

6.5.1 Accommodation problems using Head-mounted displays.

Several factors may have contributed to participants having difficulty viewing the video demonstration on the head-mounted display. A major factor may have been the individual differences in visual accommodation and how these differences would affect viewing images in a head-mounted display. A very important area of optical design that has been discussed in the ergonomics and ophthalmology literature for many years is judgement and cues of distance using display technology (Kotulak,

Morse, & Wiley, 1994; Roscoe, 1984). Recent research has explored issues concerning cues to distance that affect user's interaction with images viewed within arms' reach. The perceptual cues to space have been traditionally characterized by the presence or absence of movement or binocular information.

The effect of oculomotor (eye accommodation and convergence) responses on the perception of distance are well documented and have been the subject of much debate for several centuries (Edgar & Bex, 1995). As far as head-mounted displays are concerned binocular convergence, visual accommodation (required focus) and related reflexes have a significant influence on the personal space associated with coordination and control in interacting with these devices. Accommodation can be defined as the adjustment of the refractive power of the eye to sustain clear vision this is achieved by focusing the lens through eye movement. Vergence (convergence or divergence) is the arrangement of the lines of sight of the eyes to attain single vision from different retinal images (Kotulak & Morse, 1995). As Best, Littleton, Gramopadye, & Tyrrell (1996) point out most studies of visual activity focus on changes in regular optometric indicators (for example, visual acuity) and neglected the importance of visual accommodation. This constitutes one oculomotor response that allows viewers to see individual objects in sharp focus at different distances. Some studies for example, Fischer & Ciuffreda (1988) suggest that accommodation may be used as a depth cue for some viewers. Understanding these physiological reflexes that govern perception of depth and distance are essential as head-mounted displays are introduced into the workplace.

Our eyes work together to make viewing easy in our natural environment. Moving attention from a distant to a nearby object results in the convergence of the eyes and the accommodation of the eyes. These features of natural viewing can go askew in a head-mounted display. Visual accommodation issues with head-mounted displays alone illustrate the necessity of adjustability of focusing on a head-mounted display and the need for adjustability of the projected image in the headset. Given these considerations it was decided to switch from using a head-mounted display to a desktop P.C. for the video demonstrations of the assembly task from the second experiment onwards in the second phase of experiments.

6.6 Chapter Summary

- The second phase of experiment used a different experimental paradigm to the first phase of experiments. In the second phase of experiments the learning and the retention of an assembly task was measured. There was an initial learning phase, after which subjects were tested for their ability to build the Lego model after viewing the video clip in the conditions. There was then a criterion phase where the participants built the model without using the video clip. Finally there was a retention phase where participants came back at a later date to test their memory for the assembly. Several theorists and researchers view this paradigm as an efficient method of accessing the memory of a trained skill.
- Previous studies of learning to build model kits have used pictures and text instructions as well as filmed demonstration of the task. Most of these experiments were concerned with comparing verbal instructions (either spoken

or as text) to a combination of both. Some studies have shown an advantage for the combination conditions. Findings from of these studies reveal that the text instructions play a major role in the verbal and visual combination.

- Prior studies that manipulated media on filmed demonstrations of a procedural task are rare. Some research has indicated that filmed demonstrations with narrative instructions are not as good over time as text instructions for learning a procedure and that a filmed demonstration without narrative had the same learning outcomes as a demonstration with narrative. The conclusion from these experiments was that participants tend not to process the information to the same extent watching a narrated demonstration as when they read text instructions. However another study produced findings that showed that a filmed demonstration without any text or voice narration had the joint highest retention for a procedural task.
- There may exist a taxonomy of procedural tasks, each having characteristics that make it necessary to have different configurations of media in demonstrations. This taxonomy may be reason that would explain the conflicting results in previous studies. There are several reasons for the conflicting findings in studies investigating all formats of multimedia instructions for learning a procedural task. These contradictory findings are often caused by methodological differences between studies, such as size and quality of the visual images or amount of learner control.

- In the second phase a head-mounted display was used for the first experiment but this use was discontinued for the remainder of the experiments due to ongoing visual and physical problems with the head-mounted display. The headset was difficult to place on participants' heads and the viewing the projected images in the LCD displays may throw up several optical issues concerned with viewing projected information at such close proximity to the eyes. For the remainder of the experiments in the second phase a desktop computer was used for viewing the video clips of the assembly task.

Chapter 7

7 The Effectiveness of Different Multimedia Video Presentations for the Training and Retention of Procedural task.

“[It is] a memory profoundly differentalways bent upon action, seated in the present and looking only to the future.....”

Henri Bergson (1910) writing about what we now call procedural memory.

7.1 Chapter overview

This chapter reports the results and discusses the findings of the four experiments in the second phase of this programme of research. This phase focused on testing the efficiency of different configurations of text and voice instructions on a filmed demonstration of an assembly task. In each of the four experiments one type of display screen, desktop monitor or head-mounted display was used to view the video instructions. These experiments explored the presentation of multi modal instructions on a filmed demonstration since the latter may become a ubiquitous tool of computer-based instruction. In the four experiments participants were trained and tested in an initial phase, then tested in a criterion phase for assembling correctly a Lego model. In a retention phase they returned about a week later to be tested on their ability to build the model again in all the conditions in each experiment.

In the first part of this chapter there is an overview of the four experiments in phase two and an outline of the conditions in each. The rationale behind each experiment is discussed in turn along with the theory or theories the experiment set out to test for a procedural task. Most of the chapter will describe the design and procedure of each of

the four experiments then report the results of each experiment. At the end of the chapter there will be a general discussion of the results of this second phase of experiments.

Table 7.1: Programme of experiments for phase 2 of experiments.

Experiment	Conditions in experiment.	Theoretical Basis.
2.1	Video alone and video with vocal instructions.	Theory of multimedia learning.
2.2	Video alone, video with vocal inst., video with text inst. and video with both vocal and text inst.	Modality principle and redundancy principle.
2.3	Video alone and video with white noise.	Arousal theory.
2.4	Video with sequential vocal inst and video with simultaneous vocal inst.	Temporal contiguity principle.

Table 7.1 illustrates the programme of experiments undertaken in the second phase. In the first experiment of the research programme, experiment 2.1 there were two conditions; a video demonstration without voice instructions was compared to a video demonstration with voice instructions. This set out to test the “dual coding theory” (Paivio, 1986; Paivio, 1991) and the “multimedia principle” (Mayer & Anderson, 1992) for learning an assembly task. The contention of these theories is that learning is significantly enhanced by presenting the same learning materials in two channels, a verbal channel and a visual channel. This first experiment was carried out using a head-mounted display to display the video demonstrations.

The second experiment, experiment 2.2 had four conditions; video demonstration only, demonstration with text instructions, demonstration with voice instructions and demonstration with both text and voice instructions. This experiment tested two

theories concerning the presentation of multimedia learning. The first theory explored was the “modality principle” (Mayer & Moreno, 1998; Moreno & Mayer, 1999) . This is based on research into learning declarative knowledge with animation and the major tenet of this principle is that voice narration describing an animation is superior to an animation with descriptive text captions. The second major theory to be addressed in this experiment was the “redundancy principle” (Mayer et al., 2001). This principle contends that for multimedia learning voice narration and animation is superior to the simultaneous presentation of voice narration plus text on the animation of the same learning materials. Due to visual problems viewing the video clips on the head-mounted display reported by some participants in the experiment 2.1, a desktop P.C. was used in experiment two for viewing the video clips.

The third experiment, experiment 2.3 set out to test whether results obtained in the first and second experiments were due to an arousal effect. The advantage of the voice instructions or on screen text instructions may not have been due to information being processed semantically in the verbal and visual channels. It may simply been the case that the voice or text giving the instructions had an arousal effect on participants watching the video clip. This would enhance the participants’ recollection of where the parts were placed on the model. To test this possibility a video demonstration without text or voice instructions component was compared to a video demonstration with a burst of white noise replacing the spoken or text instructions.

The fourth and final experiment in this phase, experiment 2.4 examined the value of the “temporal contiguity principle” (Mayer & Sims, 1994; Mayer & Anderson, 1992) for learning procedural tasks. The principle suggests that a learner processes and

assimilates declarative knowledge better when the narration and the animation are presented together on a dynamic visual display. Experiment 2.4 had two conditions; in one condition the video demonstration of the model assembly was presented with the voice instruction for placing the part simultaneously with the corresponding action. In the other condition the voice instructions are heard directly before the visual action of the part being placed.

In the first and second experiments in the second phase participants who were in conditions that had text or vocal instructions were asked to fill in a question sheet where they had to put names to pictures of selected parts of the Lego model. They performed this task after the retention phase of the experiment. This test was administered to gauge how much attention the participants gave to the names given to the parts whilst they were watching the instructional video in the initial phase. This was to test the assumption made by (Baggett, 1987) that assigning names that had particular properties would aid participants' recall of the assembly task.

7.2 Experiment 2.1: The Effectiveness of a Demonstration Video with Spoken Instructions for the Learning and Retention of an Assembly Task.

A pilot was carried out as a precursor for the first experiment in the second phase and to test for any problems in the experimental paradigm that had initial, criterion and retention assembly sessions. The pilot set out to investigate possible differences between the efficiency of assembling a model either with or without voice instructions on a demonstration video. The video clips used in the pilot study showed a pair of hands assembling a model car. Four participants were trained to build the car in one of two conditions, viewing a video clip with voice instructions describing each stage of the assembly as it happened on screen or viewing a video clip without the vocal instructions. The experiment consisted of three sessions, an initial session where the participants were trained to build the model using the video clip, a criterion session where participants built the model correctly without referring to video clip and a retention session a week to ten days later when participants were asked to build the model car from memory.

There was a qualitative dimension to the pilot where the participants were given a semi-structured interview in order to gain feedback about the experiment (see appendix 7.1). The participants who were in the condition with the demonstration video with spoken instructions were also given a picture test to assess their recall of the names given to the parts of the model car used in the video demonstration with the vocal instructions. The names given to the parts were either the usual names given to car parts or reflected their physical properties. Both participants scored four points in the test. Due to the small sample of participants in the pilot study it was not possible to analyse the results in the qualitative section further. The pilot was run in

order to try the experimental paradigm with three phases; initial, criterion and retention to ascertain if there were any problems with this paradigm. Feedback from the semi-structured interviews showed that there were no reported serious difficulties with the experimental design of the pilot study. Participants found the speed of the video demonstration easy to follow; it was not too fast or too slow (see appendix 7.2). They understood the instructions and had no problems building the model car. The experimental paradigm and the procedures used in this pilot appeared to work without any problems so were not modified for the first experiment in the second phase of experiments.

This first experiment in the second phase of experiments, experiment 2.1 aimed to investigate the usefulness of “multimedia principle” and the “dual coding theory” for assembling a model from parts. The experiment also set out to investigate whether the results in the (Palmiter & Elkerton, 1993) study for the learning and retention of computer package would be the same for an assembly task. The two demonstration conditions that were used in the latter study were repeated in this experiment.

Experiment 2.1 was an extension of the pilot study with the same experimental paradigm, procedure and conditions. Participants built a Lego model in one of two conditions, either watching a video demonstration with voice instructions or without voice instructions and viewed these instructions on a head-mounted display. As with the pilot the experiment comprised of initial, criterion and retention phases. Twenty-seven participants were trained to build the same model car as used in the pilot.

Unlike the pilot this experiment had both quantitative and qualitative analyses. In experiment 2.1 the time it took participants to assemble the model and number of

times participants needed to access the video when they made a mistake or were stuck were taken as measures of how well participants had assimilated the instructions in the video clips. Participants were also asked about their views on the experiment in a semi-structured interview similar to the pilot. The participants in experiment 2.1 who were in the condition with the demonstration video with spoken instructions were also given the same picture test to as the pilot to assess their recall of the names given to the parts of the model car used in the video demonstration with the vocal instructions.

7.2.1 Method

Design

This study had a between subjects design and comprised of two parts it had a quantitative section and a qualitative one. Participants were randomly assigned to two conditions they either viewed an instructional video clip either with or without verbal instructions on how to build a Lego model of a racing car. The experimental design for the quantitative element comprised of a between participants dimension for the demonstration with voice and without voice conditions for a training to criterion session and a retention session which took place a week to ten days later.

The study had two independent variables, the between subjects independent variable was video presentation and this had two levels, whether the participant viewed the instructions with or without verbal instructions on the video clip. There were two dependent variables, the length of time taken to complete the model and the number of times the participant had to access the video clip in the training and retention phases when they did not know to do next or made a mistake. The qualitative element

of the study comprised of a semi-structured interview that elicited the participants' views about the study and a picture test was given to the participants who were in the condition with the vocal instructions on the video demonstration.

Participants

There were twenty-seven participants, twelve males and fifteen females. Thirteen were in the Video plus voice instruction condition and fourteen were in the video alone condition. Seven females and seven males were participants in the video only condition and eight females and five males were participants in the video plus voice instruction condition. The ages of the participants ranged from eighteen to thirty-seven. The participants were students and staff at the University of Abertay Dundee. All the participants were unpaid volunteers. Each participant had normal or corrected to normal eyesight.

Materials and Apparatus.

Materials in this study comprised of a computer, a Sony Glasstron binocular LCD head-mounted display that could view output from PCs. Lego bricks were used to construct a 28 piece Lego model of a car (see figure 7.1). Two video clips of a pair of hands demonstrating the stages of constructing the model, this came in two versions, one with voice instructions describing each part of the model and where it went and one without the voice narration. The sound track comprised of a voice narrative that named the parts and gave instructions where to place them (see figure 7.2). The voice narrative was relayed through speakers attached to the computer rather than the head-mounted display, this was due to the poorer sound output and difficulties in adjusting the volume for each participant on the Glasstron.

Figure 7.1: Model Car Used in Video for Pilot Study and Experiment 2.1

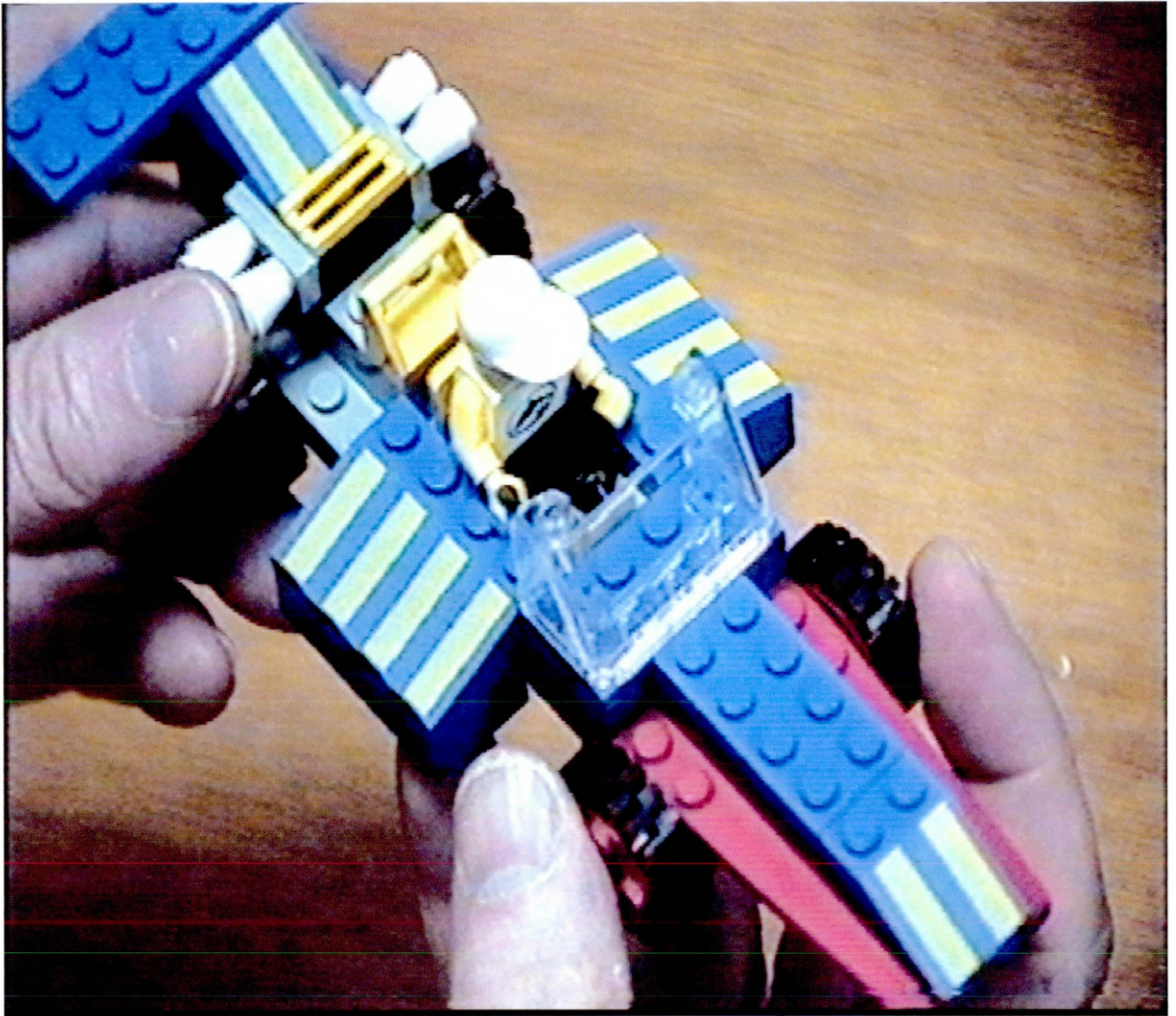


Figure 7.2: Script Used in Voice Narration for Pilot and Experiment 2.1

- 1) Bottom chassis fits underneath.....
- 2) Engine base with two spaces showing.
- 3) Front platform fits on to chassis.
- 4) Wheel axle fits on chassis front.
- 5) Wheel axle fits on the back of the top chassis.
- 6) Wheel fits on wheel axle (repeat four times).
- 7) Front bonnet fits on end of chassis front.
- 8) Body section fits on chassis front.
- 9) Body section fits across the back of chassis front.
- 10) Steering column fits on top chassis behind body section.
- 11) Seat is placed on top chassis behind steering column.
- 12) Windscreen fits on body section.
- 13) Side panel fits on top chassis left of seat.
- 14) Side panel fits on top chassis right of seat.
- 15) Body section fits beneath side panel (repeat twice).
- 16) Engine mount fits on chassis, back one space behind seat.
- 17) Head protector fits on the end of the engine mount.
- 18) Engine cowling fits on back of engine mount.
- 19) Engine grill fits on front end of engine cowling.
- 20) Exhaust fits on right side of engine cowling beneath engine grill.
- 21) Exhaust fits on left side of engine cowling beneath engine grill.
- 22) Aerofoil mount fits on engine mount with slope facing forward.
- 23) Aerofoil fits on aerofoil mount with six spaces on each side.
- 24) Driver sits in his seat.

The voice instruction for each part on the model corresponded with the visual action for placing the part. The verbal instructions were heard simultaneously in the video plus sound condition as each part was placed on the model. "Windows media player" was used to play the video clip through the head-mounted display.

Before beginning the experiment each participant was given an instruction sheet outlining the procedure of the experiment (see figure 7.3). Participants were given a list of semi-structured interview questions in experiment 2.1 that were slightly different from the pilot (see appendix 7.3). Participants in the demonstration plus vocal instruction condition were given a picture test to see how many of the names of the parts they remembered from the voice instructions (see figure 7.4). The same testing bays with recording equipment used for the first phase of experiments were used in experiment 2.1.

Procedure.

The participants were randomly assigned to one of the two conditions, video clip with voice instructions or without the voice instructions. Each participant was asked to sit in one of the testing bays that were used in the first phase of experiments (see figure 3.3a). They were then asked to read the instructions for the experiment. After they had completed reading the instructions the experimenter went over the instructions with each participant to ensure they fully understood the procedure. For analysis purposes the assembly task was divided up into three separate assemblies: (1) initial assembly, (2) criterion assembly, and (3) retention assembly. In the initial assembly session participants in both conditions sat and watched the whole of the video clip whilst wearing the Sony Glasstron head-mounted display.

Figure 7.3: Instructions for Participants in the Experiments in the Second Phase.

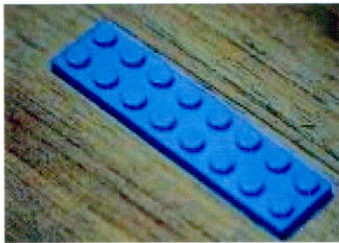
- 1) When the experimenter tells you to start, click on the play button and watch the video clip of the model being built. **Do not start building the model at this time.**
- 2) Once the video clip has finished the experimenter will give you a cue to start building the model. Try to build as much of the model as you can without referring to the instructions. If you get stuck, use the controls on windows media player to find the section of the video you are at. Once you have viewed the stage you could not remember, pause the video and continue building the model. If you make a mistake without realizing, the experimenter will tell you and you will have to look again at the part of the video you did wrong, pause the video clip and start building again. Repeat the procedure until the model is completed correctly.
- 3) The experimenter will then dismantle the model and put it back in the tray. When he gives you a cue build the model without looking at the instructions. Check with picture at the end of clip to see if you have built the model correctly. If you have not built the model correctly, study the procedure you made errors on then the model will be dismantled and you must try and build it again without looking at the instructions.

Thank you for taking part in this study.

Figure 7.4: Picture Test Given to Participants in Pilot and Experiment 2.1.

Subject

The following pictures represent parts of the model you have just constructed. In the instructional video clip you used to learn how to build the model the voice gave names to the parts of the object. Could you please write underneath the pictures what the names of the parts were. If unsure guess or leave blank. Thank you.



1 _____



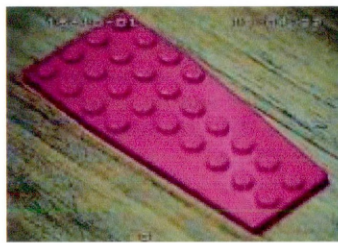
2 _____



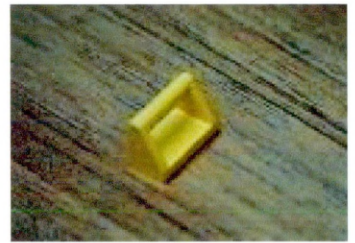
3 _____



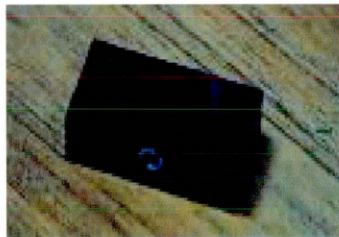
4 _____



5 _____



6 _____



7 _____



8 _____



9 _____



10 _____

Once the clip was finished the participant was asked to build the model the way shown on the video clip. They did not have to build the model in the same order as shown in the video but the parts had to go in the same places. If the participant made a mistake the experimenter who was monitoring the assembly on a nearby screen would alert the participant to the mistake, for example putting a piece in the wrong place. If the participant immediately corrected this mistake then they were permitted to continue the assembly. If however they did not know the next stage in the assembly they were advised to access the video clip to view the stage of the assembly they were having problems with. Likewise if a participant had forgotten the next stage they were advised to access the video clip. The participant used the media player to go to the stage they had problems with, viewed the stage in the video, paused the video and then went on building the model. They did this until they build the model correctly then they went on to the criterion phase.

In the criterion assembly the participant attempted to build the model without accessing the video instructions. If however the participant became stuck on a stage or stages of the construction of the model they were allowed to access the video clip then finish the assembly of the model. This assembly was not counted as the criterion assembly. Instead the participant was given a second criterion assembly trial. This second criterion assembly trial if completed successfully was counted as the criterion assembly and both the initial assembly and the first criterion trial were counted as the initial assembly phase. Upon completion of the criterion assembly participants were given an appointment to return for one more assembly seven to ten days later. In this retention assembly session built the model again but this time without looking at the video clip before the assembly. If however they had forgotten where certain parts of

the model were placed they were allowed to view the video in the same conditions as the initial assembly and build the model until it was assembled correctly. Upon completing the retention assembly the subjects were given a semi-structured interview to elicit their views about certain aspects of the study and the participants who were in the demonstration condition with the vocal instructions were given a picture test to ascertain if they remembered the names of the parts in the spoken instructions.

Because the three assembly sessions had different procedures they were analyzed separately. The procedure in the initial assembly allowed the subjects to view the entire video clip before assembly and the criterion and retention assemblies did not. In the criterion assembly the participants were not allowed to access the video clip so this assembly was analyzed separately. The retention assembly had a similar procedure as the initial assembly session, where the participants could access the video clip if they were stuck or made a mistake. However in the retention session the video clip was not viewed in its entirety before the assembly began. Each assembly had a between subjects design with two levels of the independent variable, spoken instructions or no spoken instructions. To analyze the results from the three assemblies, independent samples t-tests were carried out to test the difference in means of assembly times and frequency of video access.

7.2.2 Quantitative Results

Figure 7.5: Mean completion times and standard error in seconds in the initial training, criterion and assembly sessions for the video alone and the video with spoken instructions.

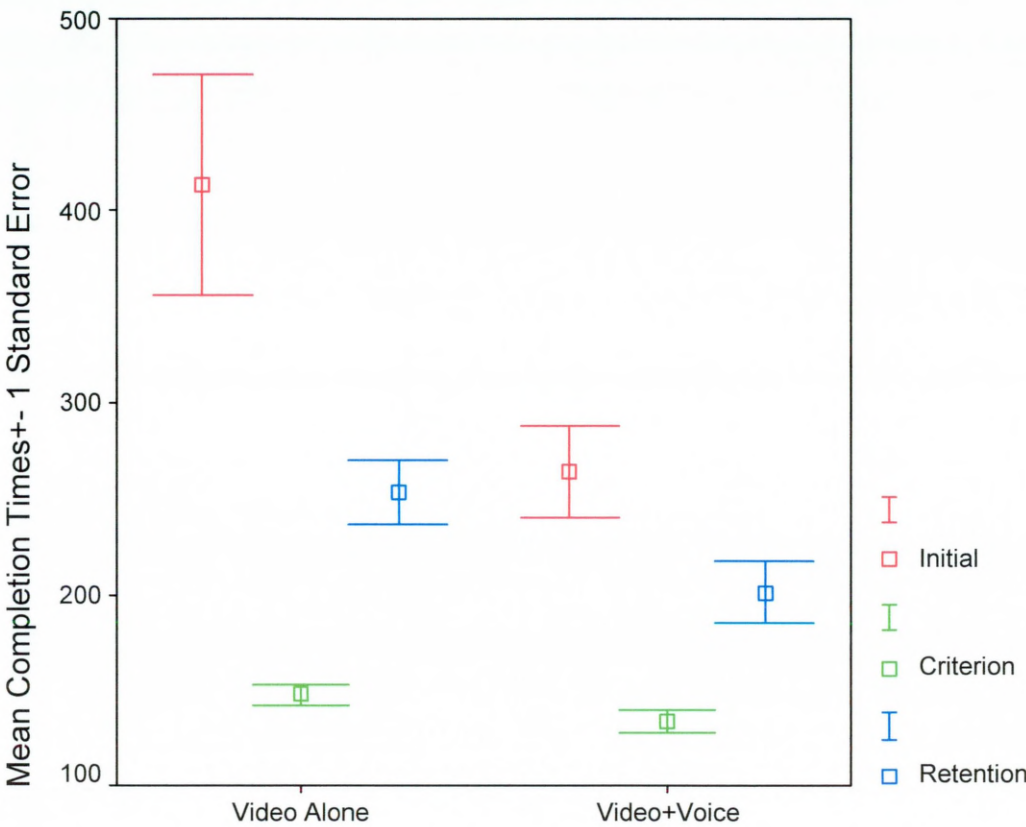


Figure 7.5 shows the mean completion times and standard error for the video alone and the video with spoken instruction conditions in all assembly sessions. For the initial assembly the video alone instruction condition had a mean completion time of 413 seconds with a standard error of 57. The video plus voice condition had a mean completion time of 264 and a standard error of 23. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was significant: $t(25)=2.404$; $p<0.05$

In the criterion assembly the video alone condition had a mean of 147 seconds, standard error of 5 seconds. The video plus sound condition had a mean of 133 seconds, standard error of 6 seconds. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was not significant: $t(25) = 1.598$; $p = \text{N.S.}$

In the retention assembly the video alone condition has a mean of 253 seconds, standard error of 16 seconds. The video plus voice condition has a mean of 200 seconds, standard error of 16. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was significant: $t(25) = 2.229$; $p < 0.05$.

Video Access Frequency

Figure 7.6: Mean number of times participants accessed the video clip in the initial and retention assembly sessions in the video only and the video plus narration conditions.

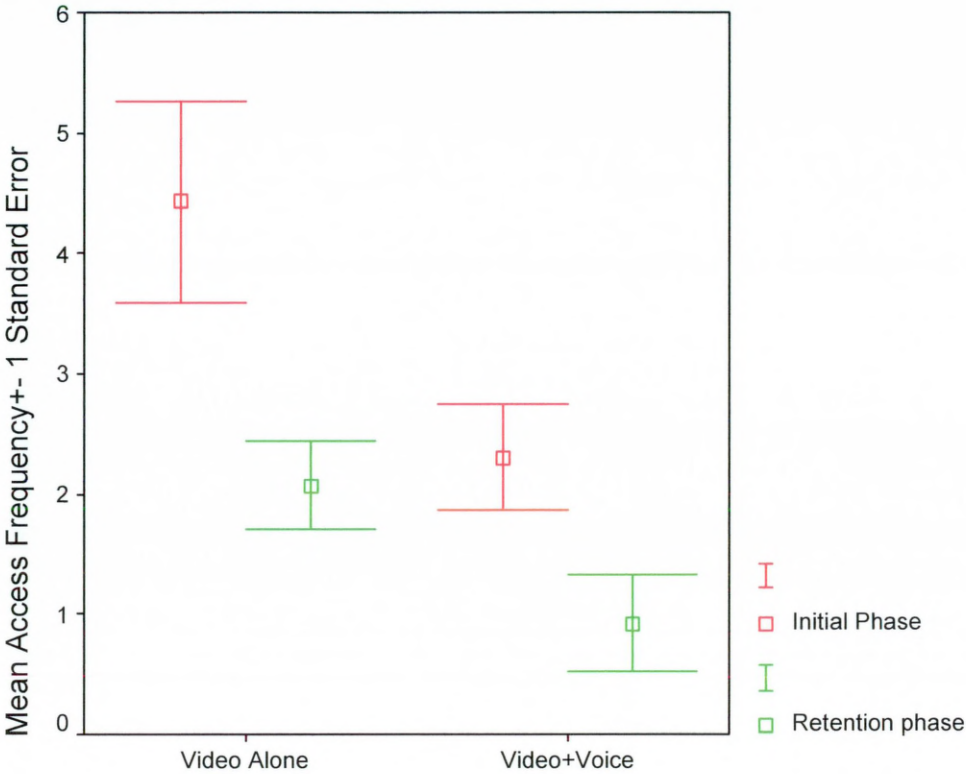


Figure 7.6: illustrates the mean number of times participants accessed the video clip when they made a mistake or were unsure of a particular stage. These measures are for the initial and retention phases for both the video alone and video with narration conditions. In the initial phase participants had a mean access frequency of 4.3, with a standard error of 0.830 for the video alone stage. Participants in the video plus voice had a mean access frequency of 2.31, with a standard error of 0.44. In the retention phase participants had a mean access frequency of 2.07, with a standard error of 0.37

for the video only condition. In this phase participants had a mean access frequency of 0.92 with a standard error of 0.4 for the video plus sound condition. To test the significance of the observable differences an independent samples t-test was carried out for the initial phase. There was a significant difference between the two conditions for mean frequency of access $t(25) = 2.535$; $p < 0.05$. An independent samples t-test was carried out for the retention phase. There is also a significant difference between the two conditions for mean frequency of access $t(25) = 2.203$; $p < 0.05$.

7.2.3 Qualitative Results.

As with the experiments in the first phase there was an interview with each participant immediately after he or she had completed the retention phase. This interview had six or seven questions about the study, depending on which condition participants were in, about the mechanics of the instruction video and the instruction formats in the two experimental conditions (see appendix 7.3). Six of the questions asked participants direct queries about the experiment. The main findings were that nearly all the participants in both conditions felt that the pace of the instructional video was just right. Only one thought the pace of the video was too slow, none thought it too fast. In question five the participants in the video only condition were asked if they would have preferred voice instructions. Three participants said they would have liked to have received verbal instructions, seven said no they would have not liked that and three said they did not know. Participants in the video plus sound condition were asked if they paid more attention to the visual information, sound information or both. Most said they paid more attention to the visual information. Another question was directed only to the participants in the video plus voice condition. The participants

were asked if they felt about having both visual and verbal information. Only two participants thought that this was too much information.

The last question was an open-ended question that asked for the participants' general comments about the experiment. Most participants found the video demonstrations of the task easy to follow and to understand. One participant thought that the film taken from the viewer's perspective made the video easier to remember. Two participants reported that the video clip was useful for giving spatial information about where each part went. A participant in the video alone condition thought that text should have been used on the video as instructions. A participant in the video plus voice instruction condition found the dual modal presentation of information very constructive. However two found the video too slow and one found building the model too easy. Several participants made comments about the usability of the head-mounted display. Two found it uncomfortable to wear a third reported feeling dizzy after using the head-mounted display and three reported problems seeing the video clip clearly.

Picture Test.

The participants in the video with voice instructions were given the same picture test as the pilot study to test their memory for the names the parts were given by the voice instructions (see figure 7.4). A correct name was awarded one point; if the participant got one of the words correct in the name they were given a half a point. For example if they called a part "engine block" instead of "engine cowling they were awarded half a point. With ten car parts in the test there was a possible maximum score of ten.

Table 7.2: Maximum and Minimum Scores plus the Mean (and standard deviations in brackets) the Correctly named parts in the Picture Test.

Minimum Score	Maximum Score	Mean (sd)
1.5	5	3.3(1.3)

As table 7.2 shows the minimum score was 1.5 and the maximum was 5. No participant could remember more than five names of the parts. The mean number of the names recalled was 3.3 with a standard deviation of 1.3.

Correct names of parts in picture test.

- 1) Aerofoil.
- 2) Body Section
- 3) Engine Mount
- 4) Exhaust
- 5) Front Platform
- 6) Head Rest
- 7) Engine Cowling
- 8) Steering Wheel
- 9)Wheel
- 10) Wheel Axle.

7.2.4 Discussion

The results of experiment 2.1 demonstrate a significant difference between the two conditions in the training and retention of the assembly task. These results indicate that participants in the video plus sound condition were significantly quicker building the model and had to refer significantly less to the instructions in both the initial training and retention phases. These results contradict the findings of the (Palmiter & Elkerton, 1993) study that found no significant differences in the efficiency of a video clip that had a voice narrative compared to one that did not have any vocal instructions for learning a computer package. The findings from the experiment

reported here demonstrate that there is an advantage for placing vocal instructions on a video demonstration for learning an assembly task. This finding supports the “the multimedia principle” (Mayer & Anderson, 1992) and the “dual coding theory” (Paivio, 1991). The findings support the idea that the verbal and visual presentation of training materials enhances learning for a procedural task as well as a declarative task. These findings may also be evidence for the taxonomy of procedural tasks outlined by (Konoske & Ellis, 1991). These findings support the possibility that different procedural tasks require different presentations of media on video or animated instructions.

Results of the post-test interviews show that a significant number of participants in the video plus vocal instruction condition said they paid more attention to the visual information than the verbal information or both. When the participants in the video alone condition were asked if they would have preferred vocal instructions only three said they could have used vocal instructions. So the perception of most of the participants in this condition was that the visual instructions were adequate for this task. The head-mounted display still posed usability problems for participants in terms of comfort and viewing the video demonstration. This feedback from the participants prompted the discontinuation of using the Glasstron in favour of completing the remainder of the experiments on the desktop computer. There was a possibility that the continued use of the head-mounted display may have a detrimental effect on the performance of some participants. The purpose of the qualitative section in this experiment was to monitor feedback from using the head-mounted display. Qualitative sections were omitted from the remainder of the experiments since they used a desktop P.C. to view the video clips.

Results from the picture test administered to the participants in the video plus vocal instruction condition show that on average less than four of the names of the parts spoken by the voice were remembered a week later. In one of the questions in the post-test interview most participants reported that they were either familiar or very familiar with the names of car parts. The participants did not process the names given to the parts very deeply whilst watching the instructional video. This contradicts the supposition by (Baggett & Ehrenfeucht, 1982) that giving the parts names that are easier for subjects to remember was an important aspect of learning to build a model kit. Experiment 2.1 indicated that participant performance was not affected by their poor recall of the names the parts were given.

7.3 Experiment 2.2: A comparison between spoken instructions, text instructions, text plus spoken instructions and no text or spoken instructions on a demonstration video for the effectiveness of learning and retention of an assembly task.

Experiment 2.2 had the same experimental paradigm as experiment 2.1; however experiment 2.2 had four experimental conditions rather than just two. Participants were asked to build a model plane in four conditions; video alone, video with voice instructions, video with text instructions and video with both text and voice instructions. Experiment 2.2 had the same procedure and experimental paradigm as experiment 2.1. Again there were three assembly sessions; initial training, criterion and retention sessions.

The rationale behind this experiment was to compare the effectiveness of the four permutations of visual and verbal instruction to determine which if any constitutes the optimum multi modal presentation of instructions for a procedural task. This experiment was constructed to test if the modality and redundancy principles (Mayer, 2000) could be used as guidelines for multimedia instructional materials. Text was used in two of the conditions to test an idea proposed by (Palmiter & Elkerton, 1993) that text instructions placed on video demonstration may be more effective for learning a procedural task than video demonstration with voice instructions. However these researchers felt that adding text instructions to a demonstration may be problematic for several reasons, for example the display may become cluttered. Baggett & Ehrenfeucht (1983) demonstrated that text instructions displayed simultaneously with the corresponding visual action displayed on a demonstration may cause difficulties for the learner since it is necessary to assimilate two pieces of

information presented in the same modality. For this reason Palmiter & Elkerton (1993) used a video demonstration with a vocal narration to impart the text instruction. The guidelines used by educational technologists for the presentation of text and corresponding section of an animation for learning declarative information is derived from work by (Mayer, Steinhoff, Bower, & Mars, 1995). These guidelines, which come under the rubric of the “temporal contiguity principle”, state that the text information should be presented simultaneously at the same time and space as the matching visual learning materials.

Michas & Berry (2000) demonstrated that the separation of text captions and the corresponding pictures of the steps of a bandaging task did not make any difference for learning a procedural task. Therefore in an attempt to circumvent the problems with presenting text simultaneously with corresponding visual instructions, conditions in experiment 2.2 with text instructions had the text instructions presented before the corresponding visual demonstration in the assembly. The voice instructions on the other hand were presented simultaneously with corresponding visual demonstration.

7.3.1 Method

Design.

This experiment had the same between subjects design as the previous experiment. However unlike the previous experiment there was no qualitative section. In experiment 2.2 participants were assigned to four conditions in the experiment. Participants watched an instructional clip either without added vocal or textual instructions, with added voice instructions, with added text instructions and with both added voice and text instructions. The experiment had the same two dependent

measures as experiment 2.1; time taken to build the model in the three phases, and the number of times the video clip had to be accessed in the initial and retention phases. In three of the conditions that had text and sound on the video clip, participants were shown pictures of parts of models and were asked to name them after completing the retention phase.

Participants.

Sixty-three participants took part in the study. Three did not return for the retention phase of the experiment, so their data could not be included in the analysis. Another participant did not obey the instructions during the experiment and therefore their data has also been omitted from the analysis, therefore data from fifty-nine subjects were used for analysis. Thirty-two of the participants were male and twenty-seven were female, they were all staff and students at Abertay University within an age range of between eighteen and forty-two. All the participants were unpaid volunteers and had normal or corrected to normal eyesight.

Materials and Apparatus.

The materials used in this experiment were a 27-piece Lego model aeroplane (see figure 7.7), a desktop computer to present the video clips on Windows Media player. There were four video demonstrations of a pair of hands building the model similar to the video demonstrations used in the previous experiment. There was a video demonstration for each condition. One without text or vocal instructions one with voice instructions, one with text instructions and one demonstration with voice and text instructions. These video demonstrations were constructed the in the same way as the previous experiment and a similar type of assembly instructions as experiment 2.1 was used in this experiment. The assembly instructions were exactly the same

whether listened to as spoken instructions or read on the screen (see figure 7.8 for script). The instruction sheet explaining the procedure of the experiment was the same as the one used in the previous experiment. The same video and analysis employed in experiment 2.1 was used in this study. There was a picture-testing sheet similar to the sheet used in the previous experiment with ten pictures of parts used in the assembly.

Procedure

The procedure in experiment 2.2 was the same as the previous experiment.

Participants were randomly allocated to the four conditions. They sat in the same booth, however in experiment 2.2 they used a desk top p.c. to view the video clip instead of the head-mounted display used in experiment 2.1.

Figure 7.7: Model of plane used in Experiment 2.2

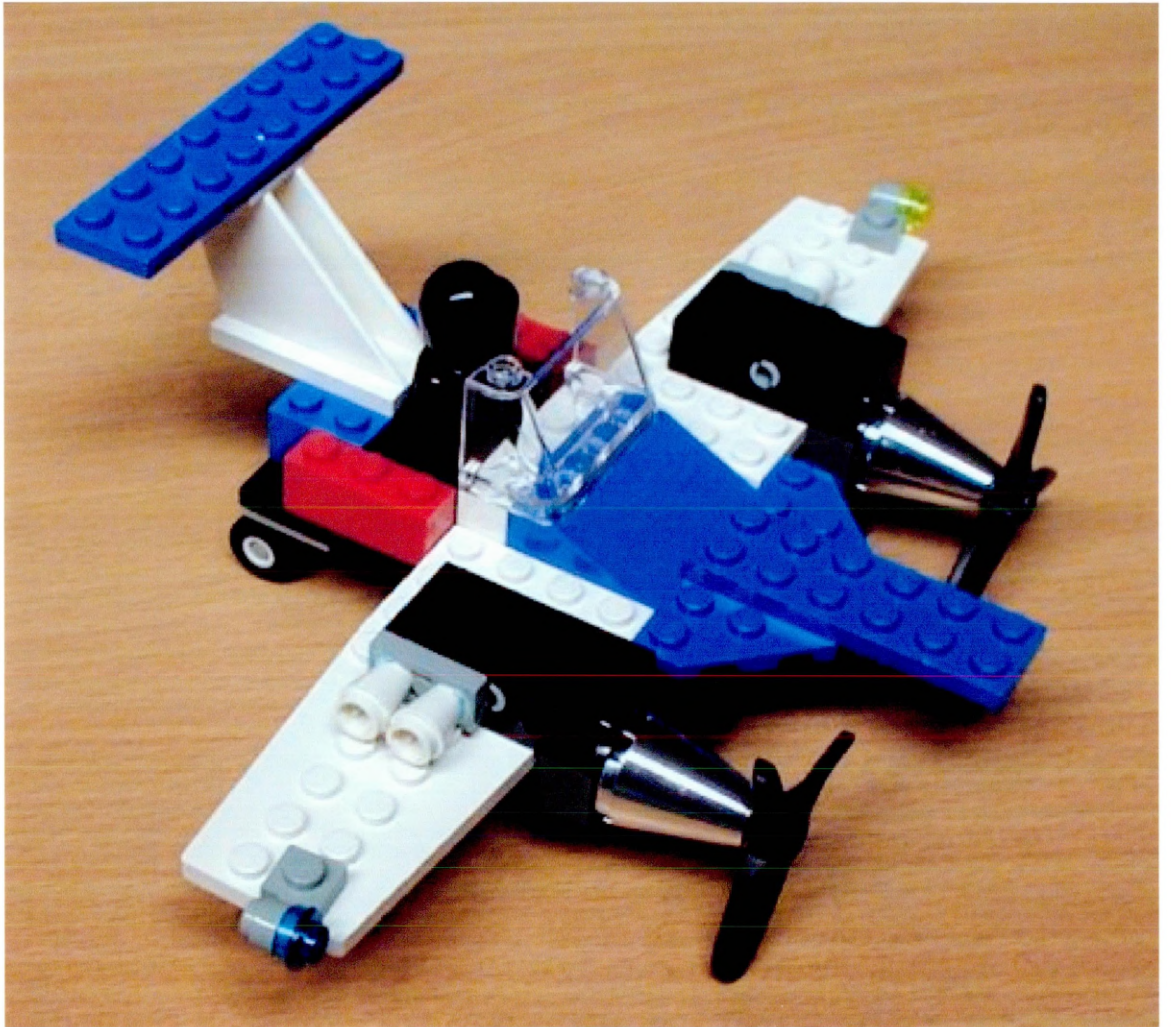


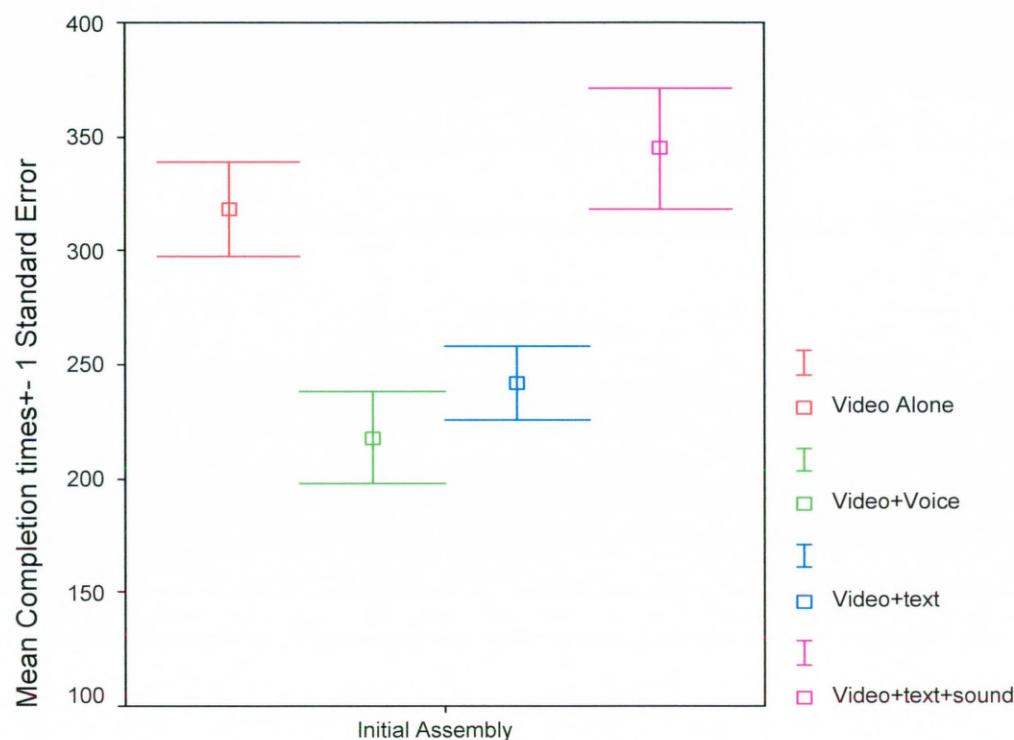
Figure 7.8: script for Experiment 2.2

- 1) Black base plate.
- 2)Wheels fit on middle front of base plate.
- 3)Wheels fit on right corner at the back of base plate.
- 4)Wheels fit on left corner at the back of base plate.
- 5)Blue triangle sits two spaces from the front of base plate.
- 6)Blue 2x6 nose section fits on middle of triangle.
- 7)Blue slope section sits behind blue triangle.
- 8)White 4x1 sits behind blue slope section.
- 9) Red 3x1 sits on right side of base plate behind white 4x1
- 10) Red 3x1 sits on left side of base plate behind white 4x1
- 11) Blue 4x2 sits at the back of the base plate.
- 12)Windscreen fits on blue slope section.
- 13)Pilot sits behind white 4x1.
- 14) White tail section fits on blue 2x4.
- 15)White tail section fits on blue 2x4.
- 16)Blue 2x8 fits on tail sections.
- 17)Right wing fits on base plate
- 18)Left wing fits on base plate.
- 19)Engine cowling sits one space forward on right wing and two spaces in from blue slope.
- 20)Engine cowling sits one space forward on left wing and two spaces in from blue slope.
- 21)Exhaust fits on back of right engine cowling.
- 22)Exhaust fits on back of left engine cowling.
- 23)Propeller shaft fits beneath front of right engine cowling.
- 24)Propeller shaft fits beneath front of left engine cowling.
- 25)Shaft guard fits on right propeller shaft.
- 26)Shaft guard fits on left propeller shaft.
- 27)Propeller fits on right shaft guard.
- 28)Propeller fits on left shaft guard.
- 29)Blue light fits forward on the end of right wing.
- 30)Yellow light fits forward on the end of left wing

7.3.2 Quantitative Results.

Initial Assembly.

Figure 7. 9: Mean Difference in Completion Times for the four Conditions in the initial assembly.



In the initial phase of the experiment participants were shown the clip before they started building the model so this phase was analysed separately from the other two phases. Figure 7.9 shows the difference in means for the completion times in seconds for the four conditions. As can be seen from this table the video plus sound plus text condition had the highest mean completion time of 336 seconds. The mean completion time for the video only condition was 321 seconds. The video plus sound had the lowest mean completion time of 222 seconds and the video plus text condition had the second lowest mean with 242 seconds. The standard error for both the video

alone and video plus voice was 19 seconds. The standard error for the video plus text condition was 16 seconds but the standard error for video plus text plus voice was the highest at 26 seconds. To further test these differences a one-way ANOVA was carried out and this was found to be significant at the one percent level, $F(3,55) = 7.602$; $p < 0.01$ (See appendix 7.4). Post hoc tests showed a significant difference between the video alone condition and both the video plus voice condition and the video plus text condition. The video plus text plus voice condition also showed a significant difference between the video plus voice condition and video plus text condition. There was no significant difference between the video alone and video plus text plus voice conditions. There was also no significant difference between the video plus voice and the video plus text conditions.

Criterion Assembly.

Figure 7. 10: Mean completion times in the criterion assembly for the four conditions.

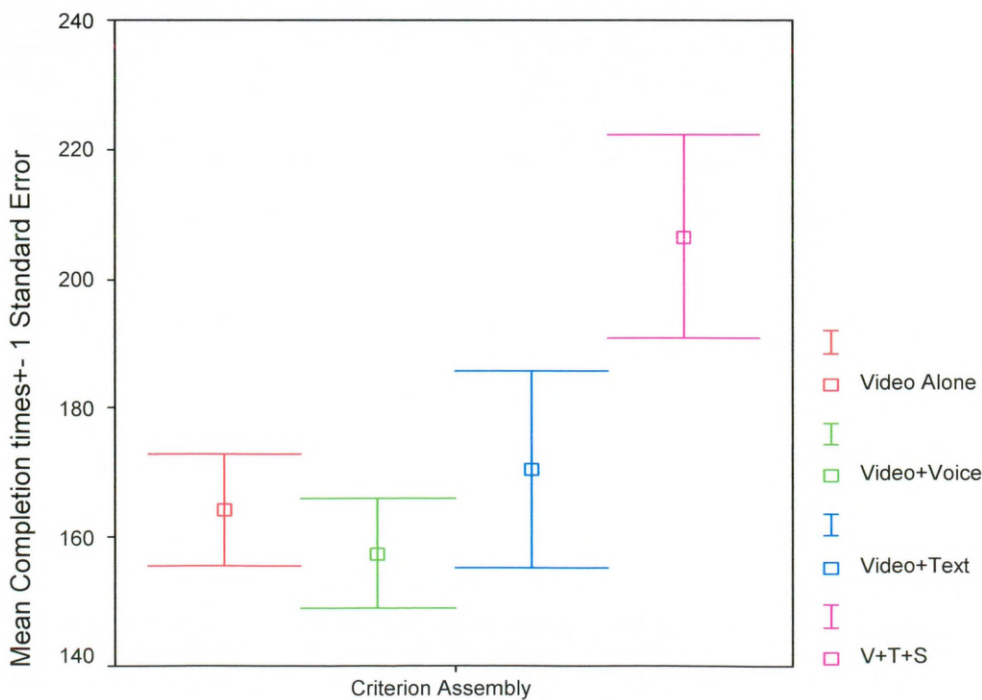


Figure 7.10 illustrates the findings from the analysis of the four conditions in the criterion assembly. The mean assembly time in the video alone condition 162.87 and the standard error was 8.12. For the video plus voice condition the mean assembly time was 159.87 and the standard error was 8.34. In the video plus text condition the mean assembly time was 170.75 with a standard error of 15.14. Finally in the video plus sound plus text condition the mean assembly time was 203 seconds with a standard error of 15.14 seconds. To test these times for significant differences this data was analysed using a one-way ANOVA. The results of this ANOVA were as follows, there was no significant effect between conditions, $F(3,55)= 2.270$; $p=N.S.$ (see appendix 7.4).

Retention Assembly.

Figure 7.11: Mean completion times for the four conditions in the retention assembly.

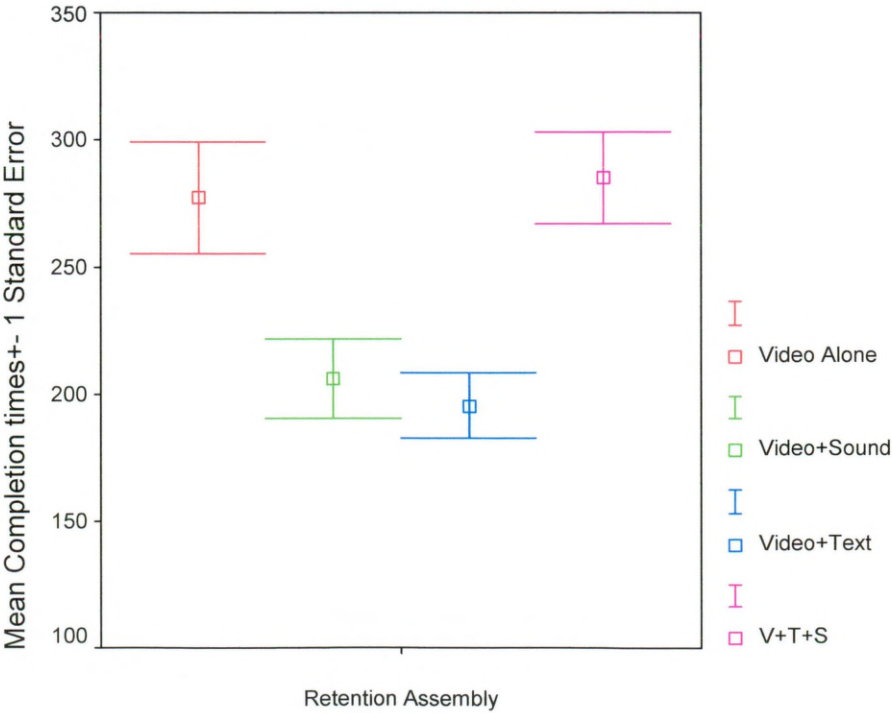


Figure 7.11 shows the mean completion times for the conditions in the retention assembly. Video plus sound plus text had the highest mean completion time with 286 seconds, video alone was the second highest with a mean completion time of 280 seconds, the video plus text had the lowest mean completion time with 196 seconds and the video plus voice condition was the second lowest with a mean completion time of 211 seconds. To further analyze these differences a one-way ANOVA was carried out on the data. There were significant differences between conditions, $F(3,55) = 7.125$; $p < 0.01$. (see appendix 7.4). Post hoc test showed a significant difference between the video alone condition and the video plus voice and video plus text conditions. The video plus text plus voice condition showed a significant difference between the video plus voice and video plus text conditions. There was no significant difference between the video and video plus text plus voice conditions. There was also no significant difference between the video plus voice and the video plus text conditions.

Figure 7.12: mean video access frequency for the four conditions in the initial assembly.

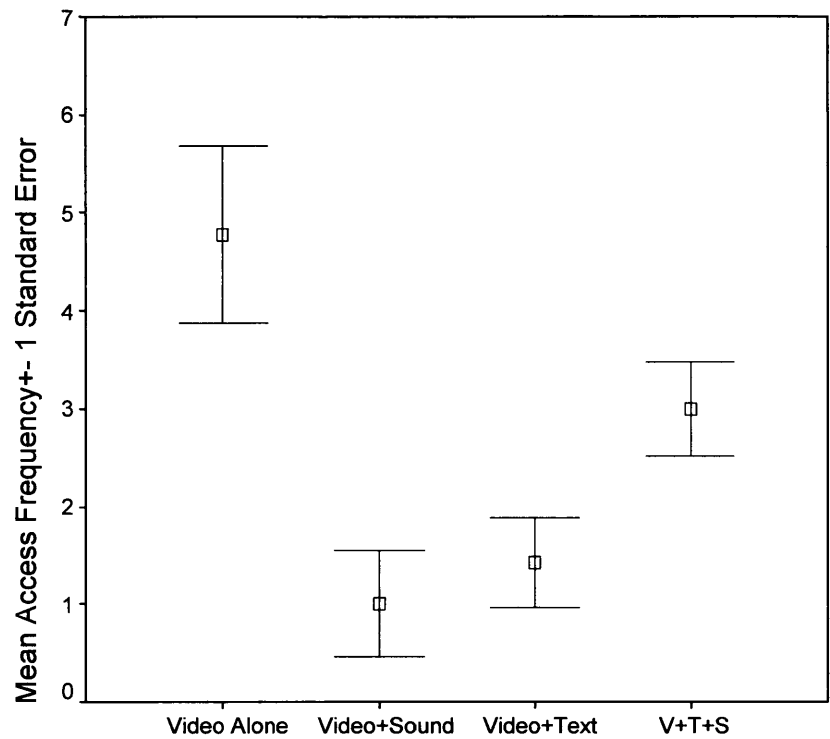


Figure 7.12 shows the mean frequency participants accessed the video clip in the initial assembly. The video alone condition had the highest mean access with 4.9 and the video and voice condition had the lowest with a mean of 1. The second lowest was the video plus text condition with 1.4 and the second highest was the video plus text plus sound condition with 2.7. To further analyse these differences a one-way ANOVA was carried out on the data and the result of this test was significant, $F(3,55)=9.124$; $p<0.01$ (see appendix 7.5). Post hoc tests indicated a significant difference between the video alone and the video plus voice and the video plus text conditions. There was a significant difference between the video plus voice and the video plus voice plus text. There was not a significant difference between the video plus text and video plus voice plus text conditions and neither was there a significant

difference between the video plus text and the video plus voice conditions. Finally there was no significant difference between the video alone and the video plus text plus voice conditions. The only difference from the mean completion times was that there was no significant difference between the video plus text plus sound and the video plus text condition.

Figure 7.13: mean video access frequency for the four conditions in the retention assembly.

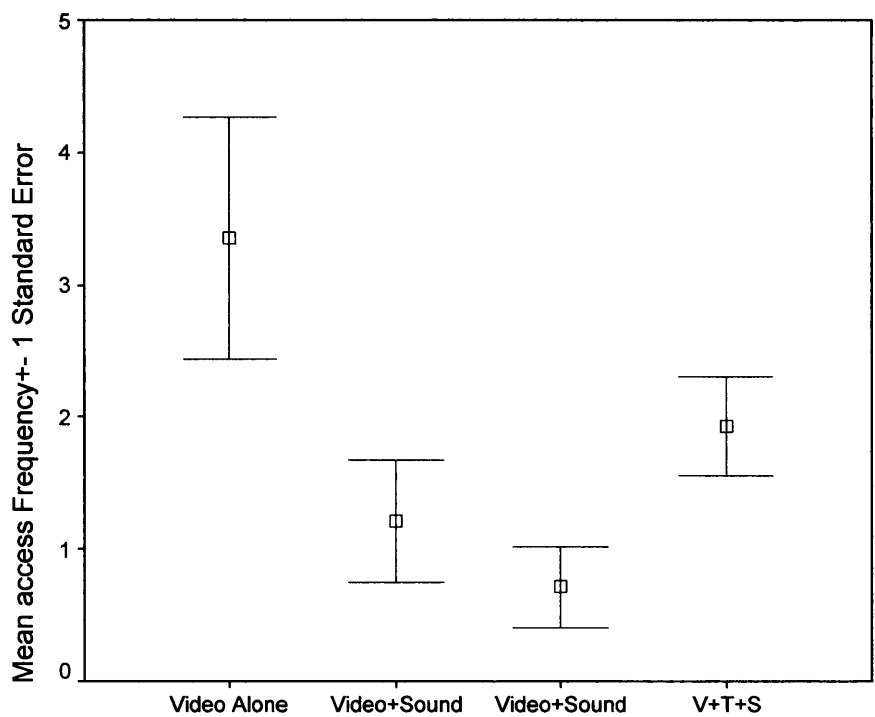


Figure 7.13 shows the mean frequency participants accessed the video clip in the retention assembly session for the four conditions. The video alone condition had the highest mean access frequency with 3.6 and the second highest mean access was the video plus text plus voice condition with 2.3. The condition with the lowest mean access was the video plus text condition with 0.7 and the second lowest was the video plus voice with 1.2. To further analyse these observable differences a one-way

ANOVA was carried out on the data and the result was significant, $F(3,55)=5.052$; $p<0.05$ (see appendix 7.5). Post hoc tests no significant difference in mean access frequency between the video plus sound, video plus text and video plus text plus sound instruction conditions. The only significant difference was between video only condition and the other conditions.

7.3.3 Picture test.

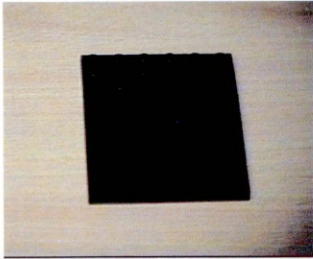
Table 7. 3:Maximum and Minimum Scores plus the Mean (and standard deviations in brackets) the Correctly named parts in the Picture Test.

Condition	Minimum Score	Maximum Score	Mean (standard deviation)
Video +Voice	0	5	2.7(1.4)
Video +Text	0	4.5	1.9(1.5)
Video +Text + Sound	0	5.5	2.1(1.7)

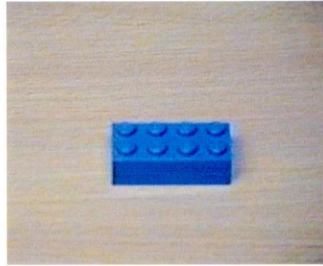
A picture test was given to participants in the conditions that had text and vocal narration on the video clip (see figure 7.14). As in experiment 2.1 participants were shown pictures of parts of models and were asked to name them after completing the retention phase. Table 7.3 illustrates the results of the picture test the maximum score was 5.5.

Figure 7.14: Picture Test Given to Participants in Experiment 2.2

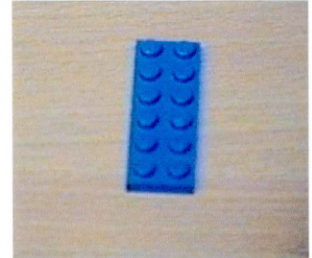
The following pictures represent parts of the model you have just constructed. In the instructional video clip you used to learn how to build the model the voice gave names to the parts of the object. Could you please write underneath the pictures what the names of the parts were. If unsure guess or leave blank. Thank you.



1



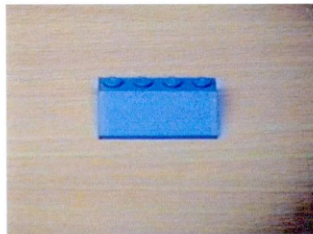
2



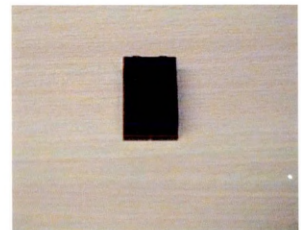
3



4



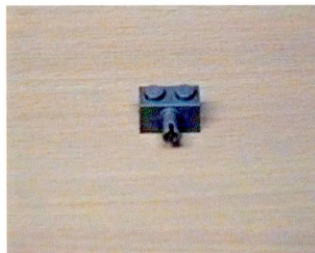
5



6



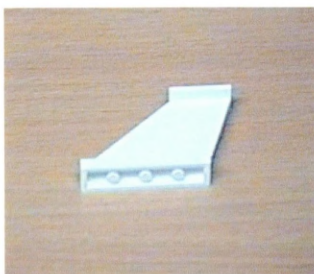
7



8



9



10

Correct names of parts in picture test.

- | | | |
|-------------------------|--------------------|---------------------------|
| 1) Black Base Plate. | 2) Blue 4x2 | 3) Blue 2x6 Nose Section. |
| 4) Blue Triangle | 5) Blue Slope | 6) Engine Cowling |
| 7) Exhaust | 8) Propeller Shaft | 9) Shaft Guard |
| 10) White Tail Section. | | |

7.3.4 Discussion

Experiment 2.2 set out to investigate whether the “modality” and the “redundancy” principles would be appropriate guidelines for multimedia visual demonstrations for learning procedural tasks. The first important finding from this experiment came from the analysis of the mean completion times in the initial phase. The results showed that the mean completion times for the video alone condition and the video plus voice plus text condition were both significantly slower than the video plus voice and video plus text conditions in the initial training assembly session. The results from the analysis of the mean completion times for the retention phase demonstrate that the significant differences between the conditions for mean completion times in the retention phase are the same as the initial phase.

Results from the analysis of the mean access frequency did show a very similar pattern graphically to the mean completion times however there were two main differences. In the initial phase mean access frequency for the video plus voice and text condition was significantly slower than the video plus voice condition but not significantly different from the video plus text condition. In the retention phase there was only a significant difference between the video alone and the other three

conditions. These latter conditions were not significantly different from one another. Results from the picture test echo the results of the previous experiment; the highest mean score was 5.5 and there were no significant differences between the conditions.

The mean completion times indicate that the optimum presentation of multi-modal video instructions for a procedural task is the visual demonstration plus some form of corresponding word based instructions, text or vocals. This result in some part contradicts the findings reported in the study by (Palmiter & Elkerton, 1993). This research revealed no difference between a demonstration with voice instructions and demonstration alone. This finding however does not support the “modality principle” (Mayer & Moreno, 1998) which states that a dynamic visual display and voice narration is superior for learning than dynamic visual display and text. Interestingly though, the relatively poorer performance of the video plus sound plus text condition compared to the video plus sound and video plus text conditions does support the “redundancy theory” outlined by Mayer et al. (2001). This heuristic maintains that text and voice narration on an animation leads to poorer learning compared to just the animation plus narration or animation plus text.

7.4 Experiment 2.3: The effect of a non verbal stimulus on a video demonstration for enhancing performance on an assembly task

The significant results in experiment 2.2 that demonstrated significant advantages for a condition with voice instructions and a condition with text instructions on a demonstration for assembling a model aeroplane. Experiment 2.3 set out to investigate whether this significant result was caused by phenomena other than the action information conveyed by the voice and text instructions. In experiments 2.1 and 2.2 the voice instruction for each stage of the assembly was heard concurrently with the visual action on the video demonstration. The sound of the voice on the video demonstration or the appearance of text may have caused an arousal effect that alerted participants and caused them to focus on the section of the video that displayed where the parts were placed on the model and possibly enhanced memory for that action.

Styles (1997) points out that human information processing is limited to the number and complexity of operations it can simultaneously perform. Moreover different conditions affect the difficulty level of the arrangement of the tasks. This difficulty may be mediated by other internal human variables, for example anxiety or fatigue. Revelle (1993) supplies a useful review of non-cognitive factors that can alter an individual's propensity to perform. A majority of these concern personality and levels of arousal. The classical theory of arousal and performance is the (Yerkes & Dodson, 1997) Law. This law states that at low level of arousal, the performance of the task is poor, as arousal increases performance increases to an optimum level. However if arousal increases past this optimum, the performance of the task goes into decline. Styles (1997) notes that situations arise where some background noise contributes to

keeping the individual alert and improves performance but if the noise becomes too loud the individual will be unable to perform any other task.

According to the arousal theory proposed by (Welford, 1968) a prolonged low event environment will adversely affect performance on a task. However any external stimulus like background noise or a sudden sound will increase arousal. In the case of the assembly task, the task performance that participants had to undertake in the second phase of experiments was the cognitive processing of video instructions. The vocal stimulus of noise or appearance of letters in the corner of the screen may have aroused the learner at a particular part of the demonstration and enhance the memory for that part of the assembly. Eysenck (1982) explains a common and simple method for studying the effects of arousal on performance is to control the level of arousal with white noise. Intense white noise has the advantage of heightening physiological arousal. Experiment 2.3 was a replication of the conditions in the two previous experiments that had demonstrations with voice instructions except the voice instructions were replaced by a burst of white noise that was heard simultaneously with the hand in the video demonstration placing a part on the model.

7.4.1 Method

Design.

Unlike the previous studies this experiment did not have a qualitative section or a picture test. The experimental design was the same as the previous experiments. Participants learned to build a Lego model of a model helicopter in two conditions. As with experiments 2.1 and 2.2 there were three phases; an initial phase, a criterion phase and retention phase. The levels of the independent variable were; white noise

on the video clip or no white noise. The dependent variables were the same as the previous two studies.

Participants

There were twenty-nine participants, thirteen males and sixteen females. Fourteen were in the video plus white noise condition and fifteen were in the video alone condition. Thirty were tested in the initial phase but one was unable to return for the retention phase. Eight females and seven males were participants in the video only condition eight females and six males were participants in the video plus white noise instruction condition. The ages of the participants ranged from nineteen to forty-four years old. The participants were students and staff at the University of Abertay Dundee. All the participants were unpaid volunteers. All the participants had normal or corrected to normal eyesight.

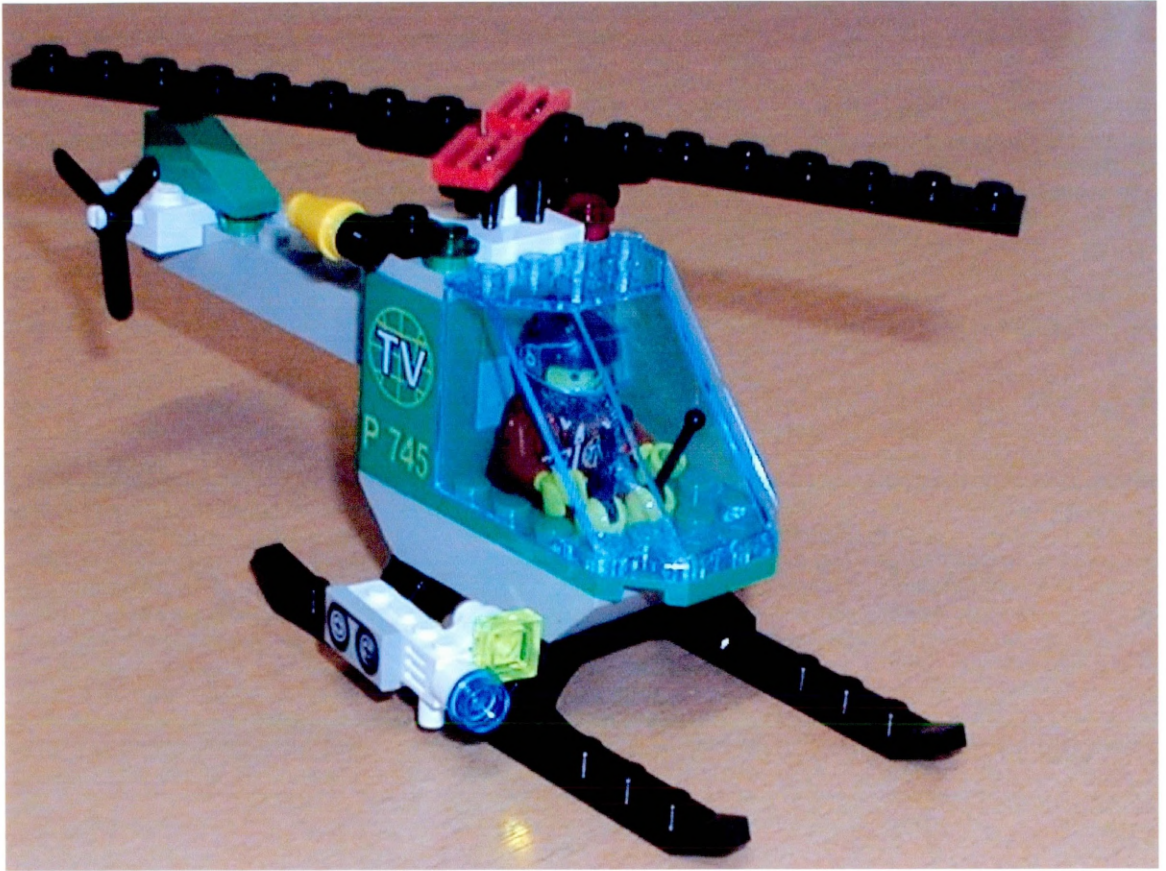
Procedure

This experiment followed the same procedure as in experiment 2.2. All participants were randomly assigned to the conditions in the experiment and shown the same instructions as in the previous experiment on a desktop computer. They learned to build the model in initial and criterion phases then returned a week to ten days later for the retention phase.

Materials and Apparatus.

The materials used in this experiment were a 30-piece Lego model helicopter (see figure 7.15) and a desktop computer to present the video clips on Windows Media player. There were two video clips; one showing a demonstration of a pair of hands building the model and the same video demonstration clip but with white noise heard

Figure 7.15: Model of helicopter used in video clip for experiment 2.3



on sections of the clip simultaneously when each part was placed on the model. The same video and analysis equipment used in experiment 2.1 and 2.2 was used in this study.

7.4.2 Quantitative Results

Figure 7.16: Mean completion times and standard error in the initial, criterion and retention assemblies for the video only and the video plus white noise conditions.

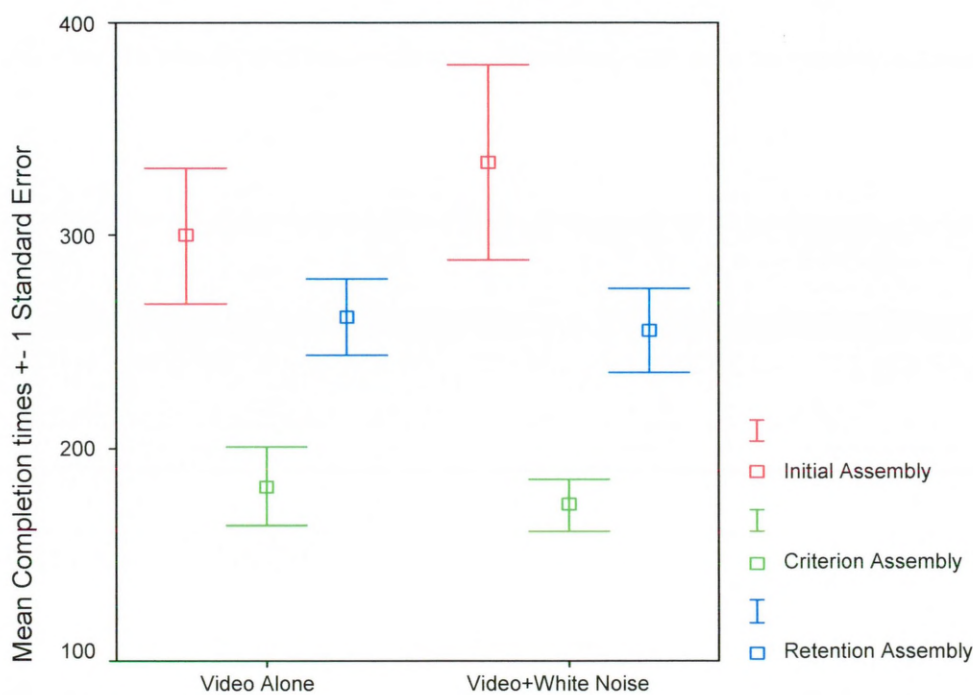


Figure 7.16 illustrates the mean completion times for the two conditions in all three assembly sessions of the experiment. In the initial phase the video alone instruction condition had a mean completion time of 299 seconds with a standard error of 31seconds. The video plus white noise condition had a mean completion time of 334 seconds and had a standard error of 45 seconds. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was not significant: $t(27)=0.628$ $p=N.S.$

In the criterion phase the video alone condition has a mean completion time of 182 seconds, standard error of 18 seconds. The video plus white noise condition has a mean completion time of 173 seconds, standard error of 12 seconds. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was not significant: $t(27)=0.385$ $p=N.S.$

In the retention phase the video alone phase has a mean completion time of 262 seconds, standard error of 17 seconds. The video plus white noise has a mean completion time of 255 seconds, standard error of 20 seconds. To test these observable differences further, the data were analysed using an independent samples t-test. The difference between means was not significant: $t(27)=0.252$ $p=N.S.$

Video Access Frequency

Figure 7.17: Mean number of times participants accessed the video clip in the Initial and Retention phases in the video only and the video plus white noise conditions.

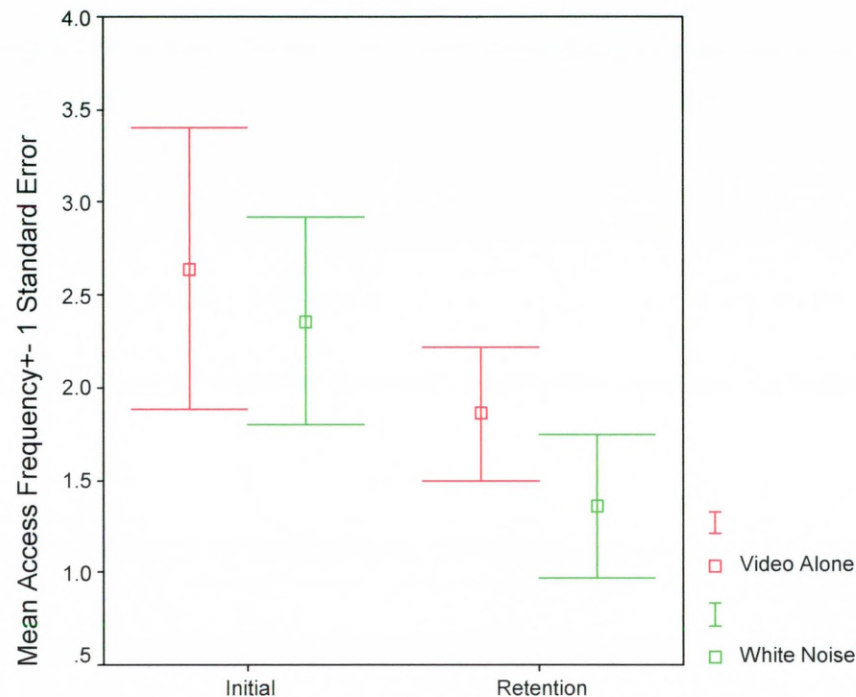


Figure 7.17 illustrates the mean number of times participants accessed the video clip when they made a mistake or were unsure of a particular stage. These measures are for the initial and retention phases for both the video alone and video with white noise conditions. In the initial phase participants had a mean access frequency of 2.6 for the video alone stage and participants in the white noise condition had a mean access frequency of 2.3. In the retention phase participants had a mean access frequency of 1.9 for the video only condition and a mean access frequency of 1.3 for the white noise condition. To test the significance of the observable differences an independent samples t-test was carried out for the initial phase. There is no significant difference between the two conditions for mean frequency of access in the initial phase $t(27)=26.023$; $p=N.S.$ An independent samples t- test was carried out for the retention

phase. There was also no significant difference between the two conditions for mean frequency of access $t(27) = 1.116$; $p = \text{N.S.}$

7.4.3 Discussion

The results experiment 2.3 indicate that white noise placed on salient points of a video demonstration that portrayed action information for each assembly stage of a Lego model did not improve the training and retention of the assembly task. Mean performances for both dependent variables did not differ significantly in the two conditions; one with white noise and one without. These findings support the idea that the semantic information in the voice and text instructions was being attended to and processed by participants in the previous experiments.

Experiment 2.3 was constructed to replicate the conditions in the two previous experiments that had demonstrations with voice instructions. However in this experiment the voice instructions were replaced by a burst of white noise that was heard at the same time with as the assembly action in the video demonstration showing a part being placed on the model. The findings of experiment 2.3 investigated the arousal theory outlined by (Welford, 1968) who maintained that an external stimulus like background noise or a sudden sound like the white noise used in the experiment will increase arousal. The results of experiment 2.3 provide evidence that participants cognitively processed the instructions and were not just aroused by their sudden appearance and later remembered that particular part of the video clip where the voice narration came in.

7.5 Experiment 2.4: An investigation into the effectiveness of the contiguous presentation of voice instruction plus video demonstration compared to the separated presentation of voice instructions plus video demonstration for an assembly task.

This last experiment in the second phase tested the “temporal contiguity principle” (Mayer, Moreno, Boire, & Vagge, 1999) for voice instructions in an assembly task.

This guideline states that learning declarative knowledge is enhanced when corresponding parts of the vocal narration and animation are presented at the same time rather than separated in time. Experiment 2.4 is an extension of the research carried out by (Michas & Berry, 2000), these researchers tested the temporal contiguity principle for learning a procedural task with illustrations and captions. These researchers found no difference between contiguously presented learning materials and those presented separately. Experiment 2.4 tested this principle on a video demonstration plus voice instructions for the training and retention of an assembly task. In one condition the vocal instructions appear simultaneously with the corresponding assembly action. In the other condition the spoken instructions end immediately before the related assembly act.

Experiment 2.4 had the same experimental paradigm as experiments 2.1, 2.2 and 2.3 and it had two experimental conditions; vocal instructions and corresponding visual information presented simultaneously and vocal instructions and corresponding visual information presented sequentially on a video clip. As with the previous experiments, experiment 2.3 had three phases, an initial training phase, the participants watched an instructional video clip on how to build a Lego model truck, a criterion phase of the experiment when the participants had to build the model again without accessing the

video clip and a retention phase when the participant returned 7-10 days later and attempted to build the model from memory.

7.5.1 Method

Design.

As with previous experiments in phase two, this experiment had a between subjects design. Participants watched an instructional video clip in one of two conditions. In one condition vocal instructions were presented simultaneously with a piece of Lego being placed on the model. In the other condition the vocal instructions were presented just prior to a piece of Lego being placed on the model. This between subjects design was used for each assembly session. The experiment had the same dependent measures as the previous experiments; time taken to build the model and the number of times the video clip had to be accessed.

Participants.

Thirty-six participants initially took part in the study. Twenty-six of the participants were male and ten were female, they were all staff and students at Abertay University Dundee. Five did not return for the retention assembly of the experiment and three neglected to follow the instructions of the study. Therefore the data of these eight subjects has also been omitted from the analysis. Data from twenty-eight subjects was used for analysis, nine females and nineteen males. The age range of the participants was between eighteen to thirty-four. All the participants were unpaid volunteers and had normal or corrected to normal eyesight.

Materials and Apparatus.

The materials used in this experiment were a 30-piece Lego model truck (see figure 7.18) and a desktop computer to present the video clips on Windows media player. There were two video demonstrations of a pair of hands building the model similar to the video demonstrations used in the previous experiment. There was a video demonstration for each condition. One with voice instructions presented just before the corresponding action information on the video demonstration and one with voice instructions presented simultaneously with the matching action information in the demonstration. These video demonstrations were constructed in the same way as the previous experiment and a similar type of assembly instructions as experiments 2.1 and 2.2 were used in this experiment (see figure 7.19). The assembly instructions were exactly the same for both conditions. The instruction sheet explaining the procedure of the experiment was the same as the one used in the previous experiment.

Procedure

This experiment followed the same procedure as the previous experiments. All participants were randomly assigned to the conditions in the experiment and shown the same instructions as in the previous experiments. They learned to build the model to criterion then returned a week to ten days later for the retention test. However there was no semi-structured interview or picture test after the retention phase.

Figure 7. 18: Model truck used in experiment 2.4.

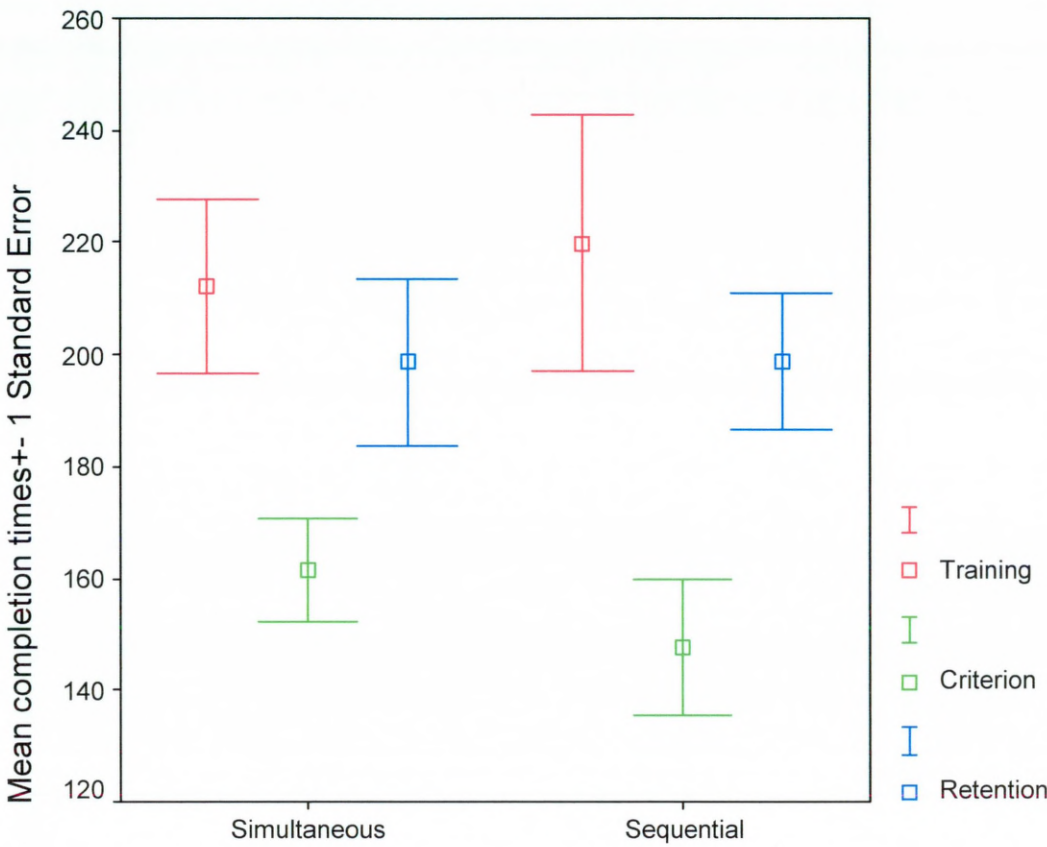


Figure 7.19: Video script for experiment 2.4

- 1) Chassis
- 2) Wheel fits on axel part of chassis
- 3) Wheel fits on axel part of chassis
- 4) Wheel fits on axel part of chassis
- 5) Wheel fits on axel part of chassis
- 6) Tow bar fits on back of chassis.
- 7) White 5 by 1 fits on right side of chassis.
- 8) White 5 by 1 fits on left side of chassis.
- 9) Engine grill sits on front of chassis.
- 10) Crane mount sits between the ends of white 5 by 1s
- 11) Crane 4 by 1 section fits on crane mount.
- 12) Crane 2 by 1 fits on crane 4 by 1 section.
- 13) Hook holder fits on crane 2 by 1 section.
- 14) Hook fits in hook holder.
- 15) Steering wheel sits behind engine grill.
- 16) Windscreen fits on top of white 4 by 1s
- 17) Motor grill sits at the back of the crane mount.
- 18) Driver sits behind steering wheel.
- 19) White 4 by 1 sits in front of motor grill.
- 20) White 4 by 1 sits on white 4 by 1.
- 21) Roof section sits on top of windscreen and white 4 by 1.
- 22) Light sits to the right on middle of roof.
- 23) Light sits to the left on middle of roof.
- 24) Side bar sits back on right of chassis.
- 25) Side bar sits back on left of chassis.
- 26) Spanner sits in front right tool holder.
- 27) Hammer sits in front left tool holder.
- 28) Flag sits left of motor grill.
- 29) Radio antenna sits on front right corner of roof section.
- 30) Metal plate fits on hook.

7.5.2 Results.

Figure 7.20: Mean difference in completion times for the voice with sequential in the initial, criterion and retention assemblies.



As can be seen from figure 7.20 in the initial assembly session the sequential condition had the highest mean completion time of 219 seconds. The mean simultaneous condition was on average 7 seconds quicker with a mean of 212 seconds. The standard error for the simultaneous voice condition was 15 seconds and the standard error for the sequential condition was higher with 22 seconds. To further test these differences an independent samples t-test was carried out and this was not found to be significant at the five percent level, $t(26) = 0.278$; $p = \text{NS}$.

In the criterion assembly the video plus sequential sound instructions had a mean completion time of 161.43 seconds. The simultaneous condition had a mean completion time of 147.1 seconds. . To further test these differences an independent samples t-test was carried out and this was not found to be significant at the five percent level, $t(26) = 0.379$; $p = \text{NS}$.

In the retention assembly session the sequential sound instructions had a mean completion time of 197.36 and the simultaneous sound instruction condition had a mean completion time of 198.51. . To further test these differences an independent samples t-test was carried out and this was not found to be significant at the five percent level. $t(26) = 0.994$; $p = \text{NS}$.

Figure 7. 21: Mean difference in access frequency for the simultaneous and sequential voice conditions in the initial assembly and the retention assembly.

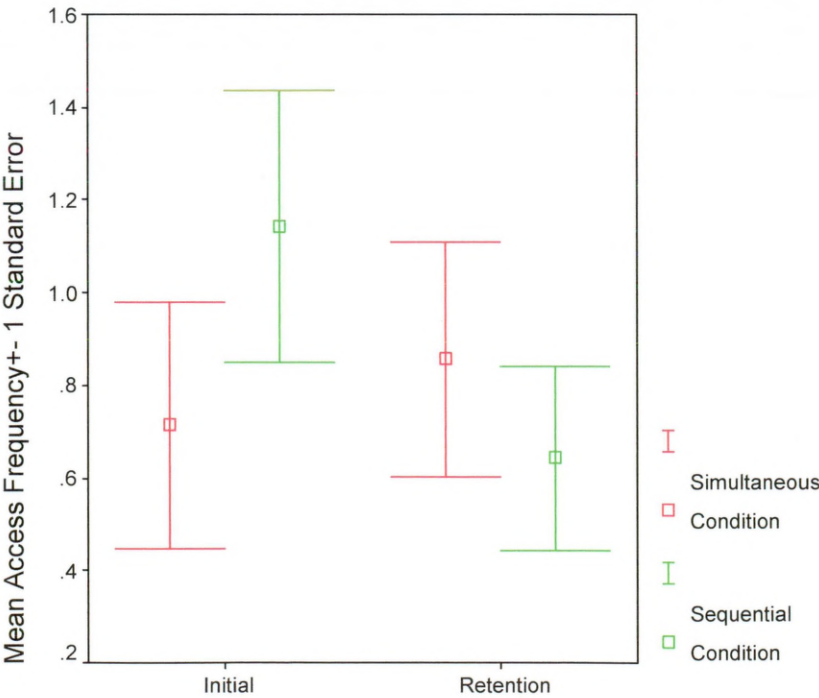


Figure 7.21 illustrates the mean number of times participants accessed the video clip when they made a mistake or were unsure of a particular stage. These measures are

for the initial and retention phases for both the video with sequential voice and video plus simultaneous voice conditions. In the initial phase participants had a mean access frequency of 0.71 for the simultaneous voice condition and participants in the sequential voice condition had a mean access frequency of 1.14. In the retention phase participants had a mean access frequency of 0.86 for the simultaneous voice condition and a mean access frequency of 0.64 for the sequential voice condition. To test the significance of the observable differences an independent samples t-test was carried out for the initial phase. There is no significant difference between the two conditions for mean frequency of access $t(26) = 0.589$; $p = \text{N.S.}$ An independent samples t-test was also carried out from the data from the retention phase. In common with the initial phase the statistical test revealed no significant difference between the two conditions, $t(26) = 0.099$; $p = \text{NS.}$

7.5.3 Discussion

Results from experiment 2.4 reveal that no significant differences between the sequential voice and the simultaneous conditions for the dependent measures. The mean completion times in the initial and retention phases were not significantly different. These results suggest that the temporal contiguity principle is not an necessary guideline for the learning and retention of an assembly task using a video demonstration. This indicates that spoken instructions and the matching visual instructions can be separated in time on a video demonstration without detracting from the performance and recall of the assembly task for the learner.

This experiment 2.4 extended the research carried out by (Michas & Berry, 2000) who tested the temporal contiguity principle for learning a procedural task with illustrations and captions. These researchers found no difference between contiguous

presented learning materials and those presented separately. Experiment 2.4 tested this principle on a video demonstration plus voice instructions for the training and retention of an assembly task. The results from this experiment challenges the use of heuristics to develop materials for learning procedural tasks based on avoiding the split attention effect (see section 5.7.1). This effect takes place when verbal and visual information are not integrated or presented simultaneously and this is postulated to cause a strain on working memory and renders the acquisition of information less effective.

7.6 General Discussion.

Table 7. 4: Summary of Quantitative results of experiments in Phase 2

Experiment	Conditions	Findings
Experiment 2.1 Opaque Glasstron Multimedia principle. Dual coding theory.	Video demonstration alone. Video demonstration plus voice instructions.	Significant difference between the conditions in both dependent variables; time and access frequency for both the initial and retention phases.
Experiment 2.2 Desktop P.C. Modality principle. Redundancy principle.	Video demonstration alone. Video demonstration plus voice instructions. Video demonstration plus text instructions. Video demonstration plus voice plus text instructions.	Significant difference between both video plus voice, video plus text conditions and video alone and video plus sound plus text conditions for completion times. Significant difference between video alone and other conditions for access frequency.
Experiment 2.3 Desktop P.C. Arousal theory.	Video demonstration alone. Video demonstration plus voice instructions.	No significant difference between the conditions in both dependent variables; time and access frequency.
Experiment 2.4 Desktop P.C. Temporal contiguity principle.	Video demonstration alone. Video demonstration plus voice instructions.	No significant difference between the conditions in both dependent variables; time and access frequency

The focus of the second phase of experiments was on ascertaining if four of the seven heuristics developed by Mayer and colleagues for learning declarative knowledge using multimedia demonstrations was the same or different for learning a procedural task. The four main theories or heuristics derived from research into declarative learning explored in phase two were; the theory of multimedia learning, the modality principle, the redundancy principle and the temporal contiguity principle.

The overall conclusion from the second phase of the experiments is that some of the heuristics developed from research on declarative learning are supported by the results for a procedural task but others are not. Some of the results from previous studies of procedural learning on video demonstrations are upheld by the results of the second phase of the experiment but on the other hand some results are incongruous with previous research. As illustrated by table 7.4, experiment 2.1 tested the multimedia principle, the findings show that this heuristic can apply to a procedural task, participants assembled the model significantly quicker and more accurately with vocal instructions on the dynamic visual display than without. Experiment 2.2 tested the second principle, the modality principle which states that vocal narration is superior for learning compared to the same information in text form on a dynamic visual display. This did not seem to have the same importance for procedural learning, text and vocal instructions were not significantly different in aiding the retention of a procedural task.

Experiment 2.2 also tested the third principle, the redundancy principle. This principle states that that both the same text and vocal information placed on a dynamic demonstration is poorer for learning than dynamic displays with just text information or vocal information. Experiment 2.2 demonstrated the redundancy principle to be a useful guideline for procedural learning since the condition with text and vocal instruction both placed on the video demonstration was shown to be significantly poorer than the two other conditions with text and narration placed on the video clip on their own. Results from experiment 2.4 demonstrated that the fourth principle, the temporal contiguity principle was not shown to be a useful guideline for procedural

learning in multimedia since the placement of the vocal narration on the video clip, whether sequential to the action information or simultaneous with this information did not significantly affect participant performance for assembling the model.

Experiment 2.3 produced evidence that learning outcomes on the video clips were not caused by a mere arousal effect. The placing of any stimulus beside the action information in the video causes observers to remember the action information better. The results of experiment 2.3 indicate that the placing of verbal information did cause the participants to remember the assembly. In two experiments, 2.1 and 2.2 participants were given a picture test to recall the names of ten parts of the model immediately after the retention phase to ascertain whether they paid attention to the names given to the parts by the text and spoken instructions. Results from these tests show that the maximum score by a participant was 5.5 and the highest mean score was 3.3. This is evidence that the participants do not process very profoundly the names given to the parts of the model.

Chapter 8

8 Synthesis of findings, general conclusions and future research.

“.....in ten years textbooks as the principle method of teaching will be as obsolete as the horse and carriage is now. I believe that visual education, the imparting of exact information through the motion picture will be a matter of course in all our schools.”

Thomas Edison (1921) cited in Large (1996).

8.1 Chapter Overview

This final chapter of the thesis will synthesize and discuss the findings of the two experimental phases of the research programme and consider their implications. The first section will review the original aims of the research and how they have developed through the two phases of experiments. Issues that were identified as important in the introductory chapters of the two phases will be returned to in this chapter. There will then be a re-examination of the aims and purposes of the first phase of experiments. After this review there will be a discussion of the conclusions drawn from this first phase. There will then be a section detailing the focus of the second phase of experiments followed by a description of the conclusions and results taken from each of the experiments in phase two. Problems encountered undertaking this research as well as alternative strategies will be discussed. There will be a section with general conclusions regarding what the implications these research findings have for using emerging technology and multimedia for learning procedural tasks. The penultimate part of the chapter explains the implications that the research findings

have for the research areas involved. The final section of the chapter describes areas of future research that would build and expand the research findings of the thesis.

8.2 Aims of research.

A major issue that this thesis sought to address concerned the usability of wearable computers for training the type of procedural task encountered in the modern workplace. The initial aim of this research programme was to investigate the optimum method of presenting multimedia-training materials for learning a procedural task using one particular emerging technology, namely the head-mounted display component of a wearable computer. This would involve finding which sub-type of head-mounted display coupled with most appropriate multimedia instructions was best suited to this task. The main usability issue with the head-mounted display was to find a subtype that was the most efficient for switching attention from the instructions displayed on the screen to the hands performing the task. A major issue with the multimedia instructions was to test whether heuristics based on research into learning a declarative task would be the same for learning a procedural task.

In phase one of the research programme the original idea was to test several types of head-mounted displays representative of models freely available on the market. The plan was to compare one of these sub-types to more conventional methods and then to one another for following instructions for a procedural task using simple linear multimedia instructions. The first phase set out to investigate three main research questions, first to ascertain if using a head-mounted display to follow procedural instructions was any different to more conventional methods of presenting information.

The second research question was to determine if one particular design of head-mounted display would be quantitatively or qualitatively more efficient for following multimedia instructions for an assembly task. This was based on the possibility that head and eye movements involved with switching between the task and the instructions would be more efficient or easier using a particular subtype of head-mounted display. The third research question addressed in this first phase was whether a head-mounted display with an opaque screen was more efficient than one with a see-through screen for following instructions to execute the assembly task. It was hypothesized that factors such as the screen position on the head-mounted display or whether the screen was see-through or opaque may have an affect on the efficient use of this component of the wearable computer.

In the first phase of experiments the initial intention was to test the efficiency and usability of several head-mounted displays, monocular or binocular and screen configurations, see-through or opaque. The instructional materials for following the assembly task were presented as animations or video clips with text or vocal instructions. After this comparison phase it was planned that the research programme move to focus on the configuration of the dynamic visual display, to determine the optimum arrangement of the media, text and vocal instructions in multimedia learning materials for learning a procedural task on the optimum design. The Multimedia instructions were to be manipulated using the head-mounted display that was demonstrated in the first phase to be qualitatively (by having the best usability preference) or quantitatively better for following assembly instructions.

8.3 Conclusions drawn from the first phase of experiments.

The general conclusions that were drawn from the first phase of experiments are that some of the original research questions could be answered. In the first experiment a head-mounted display was compared to more traditional methods of instruction; paper booklet desktop monitor or head-mounted display. The results of this experiment showed no significant differences between the two dependent variables of time taken and errors made. There was no apparent advantage for using the head-mounted display for following the instructions using a basic linear programme control in the multimedia instructions. However on the other hand there appeared to be no disadvantage for using the headset when compared to traditional instruction formats.

These quantitative findings demonstrated that a simple multimedia linear instruction programme emulating the use of a book or a manual presented on a binocular head-mounted display with an opaque screen did not differ for speed and accuracy to conventional means of conveying instructions. This outcome is different from the study by (Ockerman et al.,1997) that compared multimedia instructions based on learner control to a book for learning origami. One finding from the latter study demonstrated that participants in the wearable computer condition took significantly longer to execute the task but made fewer errors than participants in the book instruction condition. Qualitative feedback from experiment 1.1 indicated that two thirds of participants reported no difficulty in following instructions on the Glasstron head-mounted display. A proposed advantage of head-mounted displays outlined in section 2.2.1 would be that these headsets would decrease the amount and length of eye-movements between the instructions and executing the task. There was no evidence from experiment 1.1 of a speed advantage using the head-mounted display

whilst following the instructions. However practise using the head-mounted display may increase familiarity with the device, which may in turn increase performance in following instructions.

In the second experiment two head-mounted displays with different screen positions were compared for efficiency in following instructions. The results of this experiment demonstrated that participants were significantly slower and made significantly more errors when they used the Albatech HMD compared to the desktop computer condition. The participants were significantly less accurate in the Albatech condition compared to the desktop and Glasstron conditions. The qualitative feedback from the participants indicated that most preferred switching their gaze between task and instruction screen using the Glasstron compared to the Albatech. This finding provides some evidence that a binocular head-mounted display with the screen occupying a majority of the visual field is preferred to one that occupies a smaller screen area in the centre of the visual field. However the participants reported that the Albatech was more difficult to use since the information on the screen was harder to discern.

The quantitative and qualitative results in this study may have been affected not just by the position of the display screen on the head-mounted display but also by ergonomic and technological differences between the two head-mounted displays. The Albatech is not adjustable to different sizes of head and the weight of the display unit has a tendency to pull the eyeglasses forwards. To achieve a single image on the display screen the users have to line up two small mirrors by hand themselves, if this is not done correctly the users view a blurred or double image. Due to similar

technological differences a comparison between monocular and binocular head-mounted displays could not be undertaken. These results meant that phase one produced no experimental evidence for a superior head-mounted display design for following instructions. As outlined in section 2.2.2 practically all previous research into head-mounted display technology tested monocular headsets. The inability to compare different types of head-mounted displays due to technological differences and issues with adjustability meant that this research question remained unresolved. However this is an important usability issue that deserves further attention.

The only remaining research question that could be investigated was whether following the instructions are the same or different when using opaque screens compared to see-through screens on a head-mounted display. This is a pertinent research question since a number of wearable computer systems use see-through display screens. This research question could be addressed using the Sony Glasstron since it can be used in both opaque and see-through modes. The third experiment, experiment 1.3 investigated the research question involved comparing an opaque screen and a see-through screen for following assembly instructions. The results revealed no significant differences in speed and accuracy for following the instructions in the see-through mode compared to using the Glasstron in the opaque mode. The qualitative feedback from the participants indicated that they preferred using the Glasstron in opaque mode rather than the see-through mode. They found that in the see-through mode it was more difficult to switch their attention between the transparent instruction screen in the Glasstron and viewing their own hands building the Lego model. These findings of the first phase of experiments meant that the Sony Glasstron in opaque screen mode was selected to investigate the optimum

configuration of multimedia in the second phase. Of the two head-mounted mounted displays tested in the first phase, the Glasstron in opaque mode had the best qualitative feedback for usability and had the best ergonomic design.

8.4 Focus of second phase of experiments.

The first phase of experiments uncovered a number of issues and problems with comparing different types of head-mounted displays. Issues such as differences in individual eyesight may affect performance. Although the images viewed on the screens are projected to a certain distance, the display screens themselves are only a matter of millimetres away from the user's eyes. There were also different levels of adjustability in the head-mounted displays, which meant that certain types could be adjusted to fit the dimensions of a user's head better than others. The technological differences in screen display meant that particular head-mounted displays on the market have sharper and clearer displays with wider fields of view than other models. Since comparisons between head-mounted display designs could not be developed further, the second phase of experiments focused on investigating the optimum configuration of text and vocal instructions on a video clip for learning a procedural assembly. The rationale behind this second phase of experiments was to ascertain whether four of the seven heuristics or guidelines by (Mayer, 2001) for presenting text and vocal instructions on a dynamic visual display for learning declarative knowledge would be the same for learning a procedural task.

The second phase of experiments used a new experimental paradigm that comprised of three testing sessions, an initial learning phase, a learning to criterion session then a retention session approximately a week later. The Glasstron was used in the first

experiment of the second phase but was dropped in subsequent experiments due to usability issues some participants had with the headset. For the remainder of the experiments in phase two a desktop computer monitor was used to view the instructions. In the second phase participants were asked to name the parts of the models used in the video instructions in the two of the experiments to discover how deeply they processed this type of information. The idea behind this test was to investigate the technique used by (Baggett, 1987; Baggett & Ehrenfeucht, 1982) who created memorable names for all the parts of the model kit being assembled. The picture tests in the second phase set out to test the assumption that learning this type of declarative information is remembered, thus aids the recall of how the model kit was built.

8.5 Conclusions drawn from the second phase of experiments.

The second phase of experiments set out to extend the research of (Michas & Berry, 2000) into the possibility that the heuristics for the design of multimedia courseware for learning declarative knowledge were not applicable to procedural learning. The latter study manipulated mainly still pictures, line drawings and text captions. A major finding was that pictures with action information coupled with text provided the best information for learning a procedural task compared to other types of pictorial instruction formats with or without text. The second phase of experiments of this thesis moved this research to the instructional video clip. The instructional video clips used in the second phase of experiments were constructed using simple programme control. This second phase demonstrated that two of the popular heuristics created by Mayer and colleagues for learning declarative knowledge could be used for designing multimedia instructions for learning assembly tasks with video clips.

Results from the experiments in phase two provide evidence that the “multimedia principle” and the “redundancy principle” can be used for procedural tasks. The multimedia principle states that some form of voice or text instructions will give better learning outcomes when placed in the dynamic visual display than the visual demonstration on its own. This principle is derivative of the dual coding theory. The multimedia principle basically places the ideas from the dual coding theory based on pictures and text on to a dynamic visual display. The redundancy principle is the heuristic that the same text and voice instruction should not be put together on a dynamic visual display. For declarative instructions, placing text and vocal instructions together contiguously with visual information was seen to cause overload in the text and visual channel and cause a split attention effect. The text instructions to build the Lego plane appeared before the contiguous visual demonstration and vocal instruction to avoid the visual channel. It is uncertain what caused poorer performance of this condition in experiment 2.2, this finding invites further study.

However the other two heuristics that were tested did not appear to apply for procedural tasks. Experiments that tested the “modality principle” and the “temporal contiguity principle” for procedural tasks did not produce results that mirrored those for learning declarative tasks. The modality principle states that vocal narration is superior to text for learning information on a dynamic visual display. This principle is influenced by the modality effect in cognitive load theory; this states that both the auditory and visual channels must be used to relay information. Results from experiment 2.2 illustrated that this heuristic did not hold for an assembly task. There was no significant difference between the text and vocal narration conditions in the

initial training and retention phases. The third experiment, experiment 2.3 checked to see if the advantage for the verbal information on the video clips was due to an arousal effect caused by the presence of voice on the video. There was a possibility that memory for the assembly stages was enhanced by the verbal or text instructions acting as a stimulus rather than as cognitively processed instructions. Experiment 2.3 demonstrated that when white noise replaced the verbal instructions, this did not help participants to remember the assembly better than a condition with no white noise.

Experiment 2.4 tested the temporal contiguity principle and the results echo those of Michas and Berry (2000) whose findings challenged the contention by Mayer (1997) that visual and verbal instructions must be presented together since visual information must be held in working memory at the same time as corresponding verbal information. Michas & Berry (2000) produced evidence that when the temporal sequence of viewing text instructions and line drawings illustrating the procedure was different, learning outcomes were not affected. Experiment 2.4 in phase two extended this finding by showing that when the voice instructions describing each step of the assembly on a video clip are placed before the corresponding visual demonstration of the step, there is no significant difference in memory of the assembly procedure.

The results from the second phase support the findings of the Michas & Berry (2000) study that demonstrated that the heuristics for multimedia design for declarative learning are not always valid for learning procedural tasks with text and picture instructions. Such findings also indicate that the “cognitive theory of multimedia learning” conceived by Mayer and colleagues (see section 5.6) on which the four guidelines for multimedia on a dynamic visual display are based does not wholly

apply to procedural learning. The cognitive theory of multimedia learning is based on the dual coding theory, cognitive load theory and constructivism. The inconsistent findings from the experiments in phase two challenges the use of theoretical perspectives based on the cognitive load theory effect for all types of learning. The later theory underpins Mayer and colleagues' theories on presenting visual and verbal information on a dynamic visual display. In a broader sense the findings provide some evidence that supports the notion of two separate learning systems, one for declarative memory and one for procedural memory that may require different configurations of media in multimedia learning materials.

Some of the experiments in the second phase did not produce results that echoed the findings of the (Palmiter & Elkerton 1993) study. Voice narration did appear to enhance memory of a video clip of an assembly task. This result contradicts the findings of the (Palmiter & Elkerton, 1993) study that found no difference between a dynamic visual display with voice narration and one without for training participants to use a computer database. These inconsistent findings may be explained by the possibility that there is taxonomy of procedural tasks and these various tasks differ in a number of ways. Of the four types, operator task, maintenance/ repair task /assembly task, paper based task or location task, the type of task investigated by (Palmiter & Elkerton, 1991) and (Palmiter & Elkerton 1993) fits within the parameters of an operator task. However, learning to build a Lego model fits into the maintenance/ repair/ assembly category (see section 6.5 for a full description of the taxonomy of procedural tasks). Learning to use a computer programme fits the task profile associated with an "operator task". The computer task, learning to use "Hypercard Database", requires the performance on a procedure on the computer

programme that is external to the computer. The computer and computer programme being used as tools to accomplish this goal are the actual means of achieving this goal. On the other hand learning to assemble a Lego model involves performing a procedure on the component parts of an object. This assembly task requires combining the Lego parts into a complete model. The operational differences between these two tasks may mean that the configuration of the media in the multimedia learning materials has to be subtly different for each type of procedural task. It may be the case that operator tasks are better taught with visual or text only learning materials.

8.5.1 Results of the picture tests in experiments 2.1 and 2.2

In experiments 2.1 and 2.2 of the second phase participants in the conditions that had text and vocal instructions on the video clip were given an additional picture test. They were presented with a question sheet containing ten pictures of parts of the model directly after they had completed the retention phase. The participants were asked to recall the names given to the model parts in the voice narration or text instructions on the video clip. This was to assess the assumption in the study by (Baggett & Ehrenfeucht, 1982) that the names assigned to the individual parts play a role in helping participants remember the sequence of assembly when assembling a model kit. The results from the two experiments in phase two demonstrated that participants in these conditions named on average less than three parts out of ten correctly when tested. This suggests that participants did not process this information very deeply when viewing the video clip.

It is not certain from the way in which the learning was measured in the experiments in phase two whether the recall of a Lego assembly task involves explicit or implicit learning. The linking of implicit memory to procedural tasks and explicit memory to declarative knowledge has been a contentious issue in the memory literature. Some commentators, for example (Squire et al.1993) view procedural memory as a form of implicit memory whereas researchers such as (Tulving & Schacter 1990) argue that implicit memory may have similarities to both procedural and declarative memory. The above authors contend that implicit memory resembles procedural memory in that it increases perceptual skills and has similarities to declarative memory since it entails cognitive representations of the world. Berry (1992) states that explicitly learned information is information that can be verbalized and consciously recalled. The fact that the names of the parts of the Lego models constituted information that was not used to recall the assembly task supports the finding by (Michas & Berry, 2000) that participants focus on action information when learning a procedural task. This differs from declarative learning, which emphasizes the study of names and nomenclature for learning factual information.

8.6 Alternative research strategies for research.

There were several important technological differences between the Albatech and the Glasstron that may have caused the significant differences in performance between the two headsets. The Albatech can only use a television or a video signal that is inferior in clarity compared to the SVGA signal used by the Sony Glasstron. Another difference that may have contributed to a poorer performance by the Albatech is that the latter head-mounted display has a smaller field of view than the Glasstron, this smaller screen may have made the details of the Lego bricks harder to distinguish.

The implications of the above findings meant that technological and usability differences between several types of binocular and monocular head-mounted displays available for study forbade any experimental comparisons between these display types for following procedural tasks. Given these dissimilarities it would be difficult to tell if significant differences in speed and accuracy were the result of the screen position of the head-mounted display or if any significant differences are caused by technological dissimilarities such as type of signal used by the head-mounted display, field of view or by adjustability issues.

An alternative strategy for comparison studies between sub-types of head-mounted displays could have involved the development of a specially constructed head-mounted display. Such a head-mounted display could be custom built to be completely adjustable for the properties that remain fixed on various types of head-mounted displays available on the market. This custom built unit could be adjusted to project images at various distances, the screens could have adjustable fields of view and the images could be projected in different screen resolutions. The headset could have the propensity to imitate the characteristics of any head-mounted display, monocular or binocular. This would enhance greatly the possibility of evaluating screen types and screen positions without the technological differences that confound comparisons between subtypes of head-mounted display available on the market.

Another problem that emerged from all the experiments that used head-mounted displays was that the performance of participants may have been adversely affected by unfamiliarity with new technology they were using. A possible solution to this would be to provide a brief orientation session before testing that allows participants

to familiarise themselves with the visual properties of the head-mounted display. Experiments in phase one also highlighted the possibility of confounding due to participants neglecting to use eye-correction or not having the adjustments on the head-mounted set properly. To circumvent this, a battery of tests may be devised for using the head-mounted display that ensures that the technology is set up properly for the participant to use. A final issue that arose from the first phase of experiments was that the Albatech head-mounted display has a poor ergonomic design and was particularly unsuited to testing. Future research with emerging technology should include some preliminary investigation into the usability of the products that are to be used in experiments.

8.7 General conclusions and applications of research.

The practical application of the research from this thesis would be the lessons learned from the two phases of experiments. Findings from this thesis demonstrate no apparent difference between using a head-mounted display for following procedural instructions compared to traditional instruction methods when the multimedia uses a simple programme control. If a head-mounted display is the medium of delivering procedural instructions then one with an opaque screen should be considered before a head-mounted display with a transparent one. Although results did not show a significant quantitative difference between the types of screen, qualitative feedback indicated that participants were happier with an opaque screen. Designers of multimedia instructions for procedural tasks should reconsider using guidelines and heuristics based on declarative learning. Results from the second phase of experiments demonstrate that these guidelines were not always applicable to learning procedural tasks. This thesis produced evidence that present guidelines for multimedia

instructions on dynamic visual displays need not always apply to learning procedural tasks. Another guideline challenged in the second phase was that the careful naming of the parts will lead to a better memory of how the object is constructed. When instructions are being prepared for learning an object assembly, the names assigned to parts of the assembly do not have any effect on the recall of the object assembly.

The continuing use of wearable computers for training in the workplace requires research for establishing the optimum means of interaction between the user and the hardware. In a wearable computer system the screen or screens in the headset replace the computer monitor as the means of viewing information. Several subtypes of head-mounted display are currently in use: monocular displays, binocular displays, partially or full enclosed head-mounted displays. Some screens are transparent, opaque or project augmented or virtual reality images. However the differences between head-mounted displays are not confined to the characteristics mentioned above. These subtypes of head-mounted display also have different field of views, project images at different distances and have dissimilar screen resolutions. There will be in the future usability research into sub-types of other emerging technologies. This thesis has demonstrated that researchers must be made aware of the pitfalls of comparing the design of subtypes for the efficiency in performing certain tasks. Evidence from the research in this thesis indicates different versions of technology used for each subtype may make comparisons problematic. This is more marked when a cheaper product is compared to more expensive one.

The main concentration of research interest in multimedia learning at the beginning of the twenty first century is the production of learning materials for desktop computers.

This constitutes a multi-million pound industry that creates diverse products ranging from science lessons for schools, language learning for adults and training packages for industry. There is as yet no comprehensive system of guidelines or heuristics covering the different learning requirements for each knowledge domain or skill. The guidelines that do exist for configuring the optimal media presentation for computer-based learning largely reside in specialist academic journals remote from the instructional designer of multimedia products. The guidelines for multimedia learning that do exist in the mainstream publishing world are relatively rare compared to theoretical research published in journals and tend to have been developed from research into one specific learning activity.

A salient example of the restrictive nature of existing guidelines is the book on principles for creating multimedia-learning environments by (Mayer 2001). The principles described in this book are the product of many years researching the best way to teach scientific explanations using text, pictures and animations. The research that has produced the principles is founded predominately on cognitivist and constructivist theories dealing with learning concrete factual or declarative information. Research in this thesis indicated that these latter theoretical perspectives may not best be suited to learning procedural tasks. The theoretical research that appears in academic journals tends to mirror this investigation into learning one type of task by employing one theoretical perspective. Research into the cognitive load theory typifies this limited approach and likewise appears to assume that instructions for other task types will follow the resulting guidelines using predominately diagrams and text in their research.

Another problem associated with the creation of multimedia guidelines is that a great deal of current research into media presentation for learning is concerned with manipulating pictures and text. This preoccupation with more static media has resulted in the relative neglect of the important role of the dynamic visual display, whether animation or video clip in multimedia for conveying information to the learner. To illustrate this point, the research team who in recent years investigated and developed the cognitive load theory (see section 3.5) and its implications for learning are still, at the time of writing, experimenting and extending the theory using different permutations of printed text and diagrams. For example, Leahy, Chandler, & Sweller (2003) produced research that investigated the most advantageous method of explaining information on a printed graph by manipulating the presentation of text and audio accompanying the graph. Researchers working in the field of developing instructional materials should be aware of the growing importance of video technology in multimedia learning.

As far as using multimedia in tandem with emerging technologies is concerned some researchers in this particular field have made similar assumptions about the universal application of the structure of multimedia in their wearable computer system. Najjar and colleagues at the Georgia Institute of technology developed a wearable computer with a supporting multimedia system. The multimedia materials they developed for the wearable computer was an E.P.S.S. electronic performance support system (see section 1.5.4 for a full description of this system). This is a very sophisticated form of multimedia package and the above researchers assumed that it could be used to perform different types of tasks that may occur in a factory environment. The tasks tested on the wearable computer system included learning a procedural task, which

involved making origami shapes from paper (Ockerman et al., 1997b) and to enable quality control inspectors in a poultry facility to input inspection measurements into a central database from anywhere in the plant whilst inspecting products (Najjar, Thompson, & Ockerman, 1997). It may be that a wearable computer system is not suitable for performing all tasks in the work environment. The electronic performance support system used to execute these tasks is based on a complex learner control system that may not suit every task that has to be performed on the wearable system, especially for a novice learning a procedural task. Consideration must be given to the way procedural tasks are learned using a wearable computer with dynamic visual displays.

Barber & Baumann (2002) relate how the Bluetooth standard governing inter-device communication could allow different devices to share information. If a wearable computer or a similar future mobile computer-based device can be shown to be superior to an instructional manual, for training, for example access to a massive database of training situations on another remote device via emerging blue tooth technology, it remains imperative to investigate the optimum method of presenting information on such a system. Whatever nascent technology does make an appearance in the near future it is very likely that this new hardware will convey information through some sort of dynamic visual display: either an animation or a video clip. This continuing use of visual action for imparting information necessitates that the media incorporated into the dynamic visual display should be organized in the optimum way for learning or assimilating this information.

8.8 Future research.

As mentioned earlier in this chapter one way around the difficulties comparing different subtypes of head-mounted display would be to use a custom build a head-mounted display. This headset would allow the researcher to keep all the variables that constant that differ between different models on the market. This customised head-mounted display could be set to mimic all subtypes whilst keeping the same field of view, projection distance and screen resolution. An interesting finding from the second phase of experiments was that the redundancy principle was shown to be a guideline for creating multimedia on an instructional video for a procedural task. This was despite the text instructions being placed directly before the action information coupled with the narrated instructions. The reasons for this could be investigated by manipulating the distance between text instructions and simultaneous vocal and visual instructions. This might indicate a threshold where the text and vocal instructions cease interfering with one another.

It is worth reiterating the point made earlier in this thesis that research into learning procedural tasks is relatively sparse compared to research into learning declarative knowledge. The presentation of visual demonstrations of procedural tasks may continue to be incorporated into multimedia learning materials in the near future and it is important to determine the best use of media in the video clip that will improve following and retaining instructions. There has been some research over the years into the most effective way to follow instructions for object assembly. A vast majority of

the studies investigating the optimal method of presenting instructions for assembling an object have manipulated paper-base diagrams and text instructions. Some of the instructional design ideas may be incorporated into video demonstrations for following instructions for an assembly task.

In an interesting study, Novick & Morse (2000) investigated diagrams for the instruction of making paper shapes with the origami paper folding technique. In their study the first independent variable was type of instructions, text only, text with picture of finished object and step-by-step diagrams with no text. The second independent variable was the number of steps needed to complete different origami objects. The results of this study indicated that the finished picture condition was comparable in assembly accuracy to the step-by-step condition only when the number of procedural steps was small and comparable to the text only condition when the number of steps was large. This research area could be transferred to video presentations in the computer-based instruction field in an attempt to construct the optimum computer interface for following procedural instructions for assembling objects. The final picture of the assembled object could be placed on the video clip in different conditions involving size or type of object with vocal and text instructions to ascertain the best configuration for reducing working memory load and best mental model construction for the assembly task.

Findings from previous research demonstrated that text instructions require more cognitive processing therefore are remembered better over time (Walker, Jones and Mars 1983). There are also findings that text instructions are processed more thoroughly over time than a narrated demonstration (Just, 1987). Placing brief text

instructions before the corresponding visual demonstration avoids the visual channel being overloaded by the text instructions and visual demonstration. This would appear to be possible for procedural instructions since the text instructions and associated visual demonstration can be separated in time. A future experiment could involve comparing the strength of retention between a dynamic visual display with narrated instructions to one with text instructions for a procedural assembly task. It may be that text instructions placed before the visual action information will be retained better than narrated instructions.

For the enhancement of learning the procedure of a procedural task further research could expand the experimental paradigm in phase two of this thesis. Ellis et al.(1996) demonstrated that the retention of the assembly of a working model crane was enhanced significantly by using functional information in the text instructions, the latter information explains what each part does compared to a condition with structural information in the text instructions that describe spatial relationships between the model parts. The reason given for this advantage for the functional instructions was that knowing what each part does aided learners in constructing a better mental model of the assembly and in turn led to improved learning and retention of the assembly procedure. The above study used pictures of the assembly steps plus text instructions. This research into the benefits of functional information could in future studies be investigated using video clips with vocal and text instructions. The type of task tested could involve the assembly, maintenance and repair of objects. This would ascertain if this type of functional information embedded in instructions is effective in the instructional video environment.

Another important further area of research into procedural learning would be to investigate learning for other types of procedural tasks in the taxonomy of task types proposed by (Konoske & Ellis, 1991). Further studies could investigate the optimum multimedia presentation on a video clip for learning paper-work procedures and the location of objects procedural task. There could also be research into subtypes of particular procedural tasks, for example studies could be initiated to ascertain whether multimedia is the same or different for maintenance, repair or assembly tasks.

References.

Ahmad, I., Akramullah, S. M., Liou, M. I., & Kafil, M. (2001). A scalable off-line mpeg-2 video encoding scheme using a multiprocessor system. Parallel Computing, 27, 823-846.

Alty, J. L. (1993). Multimedia: We have a technology but do we have a methodology? Paper presented at the Proceedings of Ed-Media 93: World Conference on Education Multimedia and Hypermedia., Orlando, Florida.

Anderson, J. R. (1993). Problem solving and learning. American Psychologist, 48(1), 35-44.

Atkinson, R. C., & Shiffrin, R. M. (1997). Human memory: A proposed system and its control processes.

Baddeley, A. (1997). Human memory: Theory and practice revised edition: Psychology Press.

Baddeley, A. D. (Ed.). (1995). The psychology of memory: John Wiley & Sons.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. Journal of Educational Psychology, 76(3).

Baggett, P. (1986). Interactive versus passive multimedia instructions. Paper presented at the Proceedings IEEE International Conference on Systems, Man and Cybernetics.

Baggett, P. (1987). Learning a procedure from multimedia instructions: The effects of film and practice. Applied Cognitive Psychology, 1, 183-195.

Baggett, P. (Ed.). (1989). Understanding visual and verbal messages.: Elsevier Science Publishers.

Baggett, P., & Ehrenfeucht, A. (1982). How an unfamiliar thing should be called. Journal of Psycholinguistic Research., 11(5), 437-445.

Baggett, P., & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. Educational Communication and Technology Journal., 31(1), 23-32.

Bannert, M. (2002). Managing cognitive load-theory-recent trends in cognitive load theory. Learning and Instruction, 12, 139-146.

Barber, C., & Baumann, K. (2002). Embedded human computer interaction. Applied Ergonomics, 33, 273-287.

Barber, C., Haniff, D., Knight, J., & Cooper, L. (1999). Ergonomics of wearable computers. Mobile Networks and Applications, 4, 15-21.

Barron, A. E., & Atkins, D. (1993). Audio instruction in multimedia education: Is text redundancy important? Paper presented at the Proceedings of Ed-Media 95: World Conference on Educational Multimedia and Hypermedia., Graz, Austria.

Barron, A. E., & Kysilka, M. L. (1993). The effects of digital audio in computer-based training. Journal of Research on Computing in Education, 25(3), 277-289.

Bass, L., Kasabach, C., Martin, R., Siewiork, D., Smailagic, A., & Stivoric. (1997). The design of a wearable computer. Paper presented at the Conference on Human Factors and Computing Systems: CHI 97, Atlanta, Georgia.

Beccue, B., & Vila, J. (2001). The effects of adding audio instructions to a multimedia computer based training environment. Journal of Educational Multimedia and Hypermedia, 10(1), 47-67.

Becklen, R., & Cervone, D. (1983). Selective looking and the noticing of unexpected events. Memory and Cognition, 11(6), 601-608.

Berry, D. C. (Ed.). (1992). Implicit learning: Twenty-five years on. A tutorial: MIT Press.

Berry, D. C., & Dienes, Z. (1991). The relationship between implicit memory and implicit learning. British Journal of Psychology, 82, 359-373.

Best, P. S., Littleton, M. H., Gramopadye, A. K., & Tyrrell, R. A. (1996). Relations between individual differences in oculomotor resting states and visual inspection performance. Ergonomics, 39(1), 35-40.

Blake, R. (1997). Threshold for binocular rivalry. Journal of Experimental Psychology: Human Perception and Performance, 3, 251-257.

Bradey, S., & Henderson, L. (1995). Voice-overs and auditory cues: Their impact and role in learning through imm. Paper presented at the Proceedings Ed-Media 95: World Conference on Educational Multimedia and Hypermedia., Graz, Austria.

Butters, N., Wolfe, J., Martone, M., Granholm, E., & Cermak, L. S. (1985). Memory disorders associated with huntington's disease: Verbal recall, verbal recognition and procedural memory. Neuropsychologia, 23(6), 729-743.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. Educational Psychology Review, 14(1), 5-25.

Carr, C. (1992). Smart training: The managers guide to training for improved performance.: McGraw-Hill.

Carroll, J. M., Mack, R. L., Lewis, C., H., Grischkowsky, N. L., & Robertson, S. R. (1985). Exploring exploring a word processor. Human Computer Interaction., 1, 283-307.

Chandler, J., & Sweller, J. (1991). Cognitive theory and the format of instruction. Cognition and Instruction, 8(4), 293-332.

Chapman, N., & Chapman, J. (2000). Digital multimedia: John Wiley & Son.

Clark, R.E. (1983) Review of educational Research on Learning from Media.
Review of Educational Research 53(4)

Clark, T., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149-209.

Clarke, S. (1990). The crisis of fordism or the crisis of social-democracy.
Telos, 83, 71-98.

Craik, K.J.W (1952) The Nature of Explanation. Cambridge University Press.

Dick, D. (2002). The p.C. Multimedia & web handbook: Drumbeck Publishing.

Dwyer, F. M. (1982). The program of systematic evaluation: A brief review. International Journal of Instructional Media, 10, 23-38.

Edgar, G. K., & Bex, P. J. (Eds.). (1995). Vision and displays: Taylor & Francis.

Eichenbaum, H. (1997). Declarative memory: Insights from cognitive neurobiology. Annual Review of Psychology, 48, 547-572.

Ellis, J. A., Whitehall, V. W., & Irick, C. (1996). The effects of explanations and pictures on learning, retention, and transfer of an procedural assembly task. Contemporary Educational Psychology, 21, 129-148.

Ellis, S. R., & Menges, B. M. (1998). Localization of virtual objects in the near vision. Human Factors, 40(3), 415-431.

Estes, W. K. (1955). Statistical theory of distribution phenomena in learning. Psychological Review, 62(5), 369-377.

Eysenck, M. W. (1982). Attention & arousal. New York: Springer-Verlag.

Fabbro, F. (2000). Introduction to language and cerebellum. Journal of Neurolinguistics, 13, 83-94.

Farr, M. J. (1987). The long-term retention of knowledge and skills: A cognitive and instructional perspective.: Springer Verlag.

Fischer, E., Haines, R. F., & Price, T. A. (1993). Cognitive issues in head-up displays.

Fischer, S. K., & Ciuffreda, K. J. (1988). Accomodation and apperent distance. Perception, 17, 609-621.

Foyle, D. C., Mccann, R. S., Sanford, B., D.&, & Schwirzke, M. F. J. (1993). Attentional effects with superimposed symbology: Implications for head-up displays (hud). Paper presented at the Proceedings of the Human Factors and Ergonomics Society 37th meeting.

Gay, G. (1986). Interaction of learner control and prior understanding in computer assisted video instruction. Journal of Educational Psychology, 78(3), 225-227.

Gery, G. J. (1991). Electronic performance support systems: Weingarten.

Gibson, C. P. (1980). Binocular disparity and head-up displays. Human Factors, 22(4), 435-444.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. Journal of Verbal Learning and Behaviour, 5, 351-360.

Globerson, S., & Korman, A. (2001). The use of just in time training in a project environment. International Journal of Project Management, 19, 279-285.

Goldstein, E. B. (1999). Sensation and perception. (Fifth ed.): Brooks/Cole Publishing.

Gould, J. D., & Grischkowsky, N. L. (1984). Doing the same work with hard copy and with cathode ray tube (crt) computer terminals. Human Factors, 26(3), 323-337.

Gregory, R. L. (1990). Eye & brain: The psychology of seeing (Fourth ed.). London: Weidenfeld & Nicolson.

Hagman, J. D., & Rose, A. M. (1983). The retention of military tasks: A review. Human Factors, 25(2), 199-213.

Hannafin, R. D., & Rieber, L. P. (1989). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part 1. Educational Technology Research and Development, 37(2), 91-101.

Hannafin, R. D., & Sullivan, H. J. (1996). Preferences and learner control over amount of instruction. Journal of Educational Psychology, 88, 162-173.

Hay, F. J., Moscovitch, M., & Levine, B. (2002). Dissociating habit and recollection: Evidence from parkinson's disease, amnesia and focal lesion patients. Neuropsychologia, 40, 1324-1334.

Hoogeveen, M. (1997). Towards a theory of the effectiveness of multimedia systems. International journal of Human-Computer Interaction, 9(2), 151-168.

Horowitz, M. W. (1968). Organizational processes underlying differences between listening and reading as a function of complexity. The Journal of Communication, 18, 37-46.

Howarth, P. A. (1999). Oculomotor changes within virtual environments. Applied Ergonomics, 30, 59-67.

Hull, C. L. (1943). Principles of behaviour. New York: Appleton-Century-Crofts, INC.

James, W. (1890). The principals of psychology. Retrieved 19th November 2001, from <http://psychclassics.yorku.ca/James/Principles/preface.htm>

Jasper, F. (1992). The image relationship sound-image. International journal of Instructional Media, 18(2), 161-174.

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? Educational Technology Research & Development, 39(3), 5-14.

Just, M. A., & Carpenter, P. A. (1987). The psychology of reading and language comprehension.: Allyn & Bacon.

Kalyuga, S., Chandler, J., & Sweller, J. (1998). Levels of expertise and instructional design. Human Factors, 40(1), 1-17.

Kester, E., Kirschner, P. A., Jeroen, J. G., Van Merriënboer, J. J. G., & Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills., 17, 373-391.

Kettanurak, V., Ramamurthy, K., & Haseman, W. D. (2001). User attitude as a mediator of learning performance improvement in an interactive multimedia environment:

An empirical investigation of the degree of interactivity and learning styles. International journal of Human-Computer Studies, 54, 541-583.

Kinzie, M. B. (1990). Requirements and benefits of effective interactive instruction: Learner control, self regulation, and continuing motivation. Educational Technology Research & Development, 38(1), 1-21.

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. learning and Instruction, 12, 1-10.

Kirsh, D. (1995). The intelligent use of space. Artificial Intelligence, 73, 31-68.

Konoske, P. J., & Ellis, J. A. (Eds.). (1991). Cognitive factors in learning and retention of procedural tasks. New York: Praeger.

Kotulak, J. C., Morse, D. G., & Wiley, R. W. (1994). The effect of knowledge of object distance on accommodation during instrument viewing. Perception, 23, 671-679.

Kotulak, J. C., & Morse, S. E. (1995). Oculomotor responses with aiviator helmet mounted displays and their relation to in-flight symptoms. Human Factors, 37(4), 699-710.

Krakow, G. (2002). Here come tablet pcs, Retrieved 13th September 2002 <http://www.msnbc.com/news/771540.asp?cp1=1>.

Laramée, R. S., & Ware, C. (2002). Rivalry & interference with a head-mounted display. ACM Transactions on Human Computer Interaction, 9(3), 238-251.

Large, A. (1996). Computer animation in an instructional environment. Library Information & Research, 18, 3-23.

Large, A., Behesti, J., Breuleaux, A., & Renaud, A. (1994). Multimedia and comprehension: A cognitive study. Journal for the American Society for Information Science., 46, 340-347.

Large, A., Besheti, J. B., Breleux, A., & Renaud, A. (1995). Multimedia and comprehension: The relationship among text, animation and captions. Journal of the American Society for Informational Science, 46(5), 340-247.

Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. Instructional Science, 25, 117-131.

Leahy, W., Chandler, J., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. Applied Cognitive Psychology., 17(4), 401-418.

Lee, W. W., & Owens, D. L. (2000). Multimedia-based instructional design. San Fransisco.: Jossey-Bass/Pfeiffer.

Leigh, R. J., & Zee, D. S. (1999). The neurology of head-movements. (Third ed.): Oxford University Press.

Levie, H. W., & Lentz, R. (1982). Effect of text illustrations: A review of research. Education Communication and Technology Journal, 30(4), 195-232.

Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. Cognitive Psychology, 17, 26-65.

Lin, R., & Kreifeldt, J. G. (2001). Ergonomics in wearable computer design. International Journal of Industrial Ergonomics, 27, 259-269.

Mack, R. L., Lewis, C., H., & Carroll, J. M. (1983). Learning to use word processors: Problems and prospects. A.C.M. Transactions on Office Information Systems, 1, 254-271.

Mann, S. (1997) Wearable Computing: A First Step Toward Personal Imaging. Cybersquare Computer, Vol. 30, No. 2, February

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. Journal of Educational Psychology, 81(2), 240-246.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 81, 1-19.

Mayer, R. E. (2001). Multi-media learning: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. Journal of Educational Psychology, 83, 484-490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. Journal of Educational Psychology, 84(4), 444-452.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. Journal of Educational Psychology, 88(1), 64-73.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words. Journal of Educational Psychology, 82(4), 715-726.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. Journal of Educational Psychology, 93(1), 187-198.

Mayer, R. E., & Moreno, R. (1998). A split attention effect in multimedia learning: Evidence for dual processing in working memory. Journal of Educational Psychology, 90(2), 312-320.

Mayer, R. E., & Moreno, R. (2002a). Aids to computer-based multimedia learning. Learning and Instruction, 12(1), 107-119.

Mayer, R. E., & Moreno, R. (2002b). Animation as an aid to multimedia learning. Educational Psychology Review, 14(1), 87-99.

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. Journal of Educational Psychology, 91(4), 638-643.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. Journal of Educational Psychology, 86(2), 389-401.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design:

Using annotated illustrations to foster meaningful learning of science text. Educational Technology Research and Development, 43(1), 31-43.

Melzer, J. E., & Moffit, K. (1997). Head-mounted displays. Designing for the user: McGraw Hill.

Michas, I. C., & Berry, D. C. (2000). Learning a procedure task: Effectiveness of multimedia presentations. Applied Cognitive Psychology, 14, 555-575.

Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. The Psychological Review, 63(2), 81-97.

Millodot, M. (2000). Dictionary of optometry and visual science (Fifth ed.): Butterworth Heinemann.

Molitor, S., Ballstaedt, S., & Mandle, H. (Eds.). (1989). Problems in knowledge acquisition from text and pictures.: Elsevier Science Publishers.

Moreno, R. M., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. Journal of Educational Psychology., 91(2), 358-368.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87(2), 319-334.

Najjar, L. J. (1995a). A review of the fundamental effects of multimedia information presentation on learning, Technical Report GIT-GVU-95-20: Georgia Institute of Technology.

Najjar, L. J., Thompson, C., & Ockerman, J. J. (1999). Using a wearable computer for continuous learning and support. Mobile Networks and Applications, 4, 69-74.

Najjar, L. J., Thompson, J. C., & Ockerman, J. J. (1997). A wearable computer for quality assurance inspectors in a food processing plant. Paper presented at the Digest of papers: First International Symposium on Wearable Computers., Los Alamitos, CA.

Newell, A., & Simon, H. A. (1972). Human problem solving. New Jersey: Prentice-Hall.

Newman, W. S., Podgurski, A., Quinn, F. L., Merat, F. L., Branicky, M. S., Barendt, N. A., Causy, G. C., Hassler, E. L., Kim, Y., Swaminathan, J., & Velasco, V. (2000). Design lessons for building agile manufacturing systems. IEEE Transactions on Robotics and Automation, 16(3), 228-238.

Novick, L. R., & Morse, D. G. (2000). Folding a fish, making a mushroom: The role of diagrams in executing assembly procedures. Memory and Cognition, 28(7), 1242-1256.

Ockerman, J. J., Najjar, L. J., & Thompson, D. (Eds.). (1997a). Evaluation of a wearable computer performance support system. Charlottesville: AACE.

Ockerman, J. J., Najjar, L. J., & Thompson, D. (1999). Fast: Future technology for today's industry. Computers in Industry, 38, 53-64.

Ockerman, J. J., Najjar, L. J., & Thompson, J. C. (1997b). Wearable computers for performance support: Initial feasibility study. Paper presented at the First International Symposium on Wearable Computers, Los Alamitos California.

Ockerman, J. J., Najjar, L. J., Thompson, J. C., Treanor, C. J., & Atkinson, F. D. (1996). Fast: A research paradigm for educational performance support systems.

Paper presented at the Educational Multimedia and Hypermedia 1996, Charlottesville, Virginia.

Paivio, A. (Ed.). (1983). The empirical case for dual coding: Lawrence Erlbaum Associates.

Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255-287.

Paivio, A., & Caspo, K. (1973). Picture superiority in free recall: Imagery or dual coding. Cognitive Psychology, 5, 176-206.

Paivio, A., Clarke, J. M., & Khan, M. (1988). Effects of concreteness on composite imagery rating and cued recall. Memory and Cognition, 16(3), 422-430.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory. Journal of Experimental Psychology, 20(5), 1196-1204.

Palmiter, S., And Elkerton, J., & Baggett, P. (1991). Animated demonstrations vs written instructions for learning procedural tasks: A preliminary investigation. International Journal of Man Machine Studies, 34, 687-701.

Palmiter, S., & Elkerton, J. (1991). An evaluation of animated demonstrations for learning computer-based tasks. Paper presented at the Conference on Human Factors and Computing Systems. CHI 91, New Orleans, U.S.A.

Palmiter, S. A., & Elkerton, J. (1993). Animated demonstrations for learning computer based tasks. Human Computer Interaction., 8, 193-216.

Park, O. A., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays: A review. Instructional Science, 21, 427-449.

Parkin, A. J. (1993). Memory phenomena, experiment and theory: Blackwell Publishers.

Parkin, A. J. (Ed.). (2001). The structure and mechanisms of memory. Philadelphia: U.S. Psychology Press.

Peek, J. (Ed.). (1989). Trends in the delayed use of information from an illustrated text.: Elsevier Science Publishers.

Peli, E. (2001). Optometric and perceptual issues with head-mounted display (hmd).

Piantanida, T. (1993). Another look at hmd safety. Cyberedge Journal, Nov/Dec, 9-12.

Poire, M. J., & Sable, C. F. (1984). The second industrial divide. New York: Basic Books.

Pollock, E., Chandler, J., & Sweller, J. (2002). Assimilating complex information. learning and Instruction, 12, 61-86.

Preece, J. (1993). Hypermedia, multimedia and human factors. In J. Williamson & C. R. Latchen (Eds.), Interactive multimedia: Practice and promise. (pp. 135-150): Kogan Page.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. Psychological Bulletin, 88(1), 16-45.

Reber, A. S. (1993). Implicit learning and tacit knowledge: An essay on the cognitive unconscious.: Oxford University press.

Reimann, P. (2003). Multimedia learning: Beyond modality. learning and Instruction, 13, 245-252.

Revelle, W. (Ed.). (1993). Individual differences in personality and motivation: "non-cognitive" determinants of cognitive performance. Oxford: Oxford University Press.

Ritzer, G. (1992). Sociological theory (Third ed.): McGraw Hill.

- Rohaly, A. M., & Karsh, R. (Eds.). (1999). Helmet mounted displays.: Research Studies Press Limited.
- Roscoe, S. N. (1984). Judgements of size and distance with imaging displays. Human Factors, 26(6), 617-629.
- Roscoe, S. N. (Ed.). (1993). The eyes prefer real images.: Taylor and Francis.
- Royer, J. M. (1979). Theories of the transfer of learning. Educational Psychologist, 14, 53-69.
- Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. Psychological Bulletin, 95(3), 355-386.
- Sanford, B. D., Foyle, D. C., Mccann, R. S., & Jordan, K. (1993). Head-up displays: Effect of information location on the processing of superimposed symbology. Paper presented at the Seventh International Symposium on Aviation Psychology., Ohio State University.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science, 3(4), 207-217.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human processing: 1. Detection, search, and attention. Psychological Review, 84(1), 1-25.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. Learning and Instruction, 13, 141-156.

Scoville, B. S., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery & Psychiatry, 20(11).

Shahrokh, M. C., R. (1998). Agile manufacturing-an enabling technology for competeing in changing markets. Semi-Conductor Fabtech, 8, 27-32.

Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. Computers & Education, 34, 177-193.

Shyu, H. S., & Brown, S. W. (1995). Learner-control: The effect of learning a procedural task during computer-based video instruction. International journal of Instructional Media, 22(3), 85-96.

Skinner, B. F. (1948). 'superstition' in the pidgeon. Journal of Experimental Psychology, 38, 168-172.

Smith, J., Siegert, R., & Mcdowall. (2001). Preserved implicit learning on both the serial reaction time task and artificial grammer in patients with parkinson's disease. Brain and Cognition, 45, 378-391.

Sojourner, R. J., & Antin, J. F. (1990). The effects of a simulated head-up display speedometer on perceptual task performance. Human Factors, 32(3), 329-339.

Spangenberg, R. W. (1973). The motion variable in procedural learning. A.V. Communication Journal, 21, 419-436.

Squire, L. R., & Cohen, N. J. (1980). Preserved learning and retention of pattern-analysis skill in amnesia. Science.

Squire, L. R., & Kandel, E. R. (1999). Memory: From mind to molecules. New york: Scientific American Library.

Squire, L. R., Knowlton, D. S., & Musen, G. (1993). The structure and organisation of memory. Annual Review of Psychology, 44, 453-459.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe system. Science., 253, 1381-1386.

Squire, L. R., & Zola-Morgan, S. (1996). Structure and function of declarative and nondeclarative memory systems. Paper presented at the Memory: Recording Experience in Cells and Circuits Colloquium., National Academy of Sciences, Irvine, C.A.

Stone, D. E., & Glock, M. D. (1981). How do young adults read directions with and without pictures. Journal of Educational Psychology, 73(3), 419-426.

Styles, E. A. (1997). The psychology of attention: Psychology Press.

Sweller, J. (1988). Cognitive load during problem solving. Cognitive Science, 12, 257-285.

Tenenbaum, G., Naidu S., Jegede, O., & Austin, J. (2001). Constructivist pedagogy in conventional on-campus and distance learning practice: An exploratory investigation. Learning and Instruction, 11, 87-111.

Thompson, J. C., Najjar, L. J., & Ockerman, J. J. (1997). Wearable computer based training and performance support systems. Paper presented at the 19th Interservice/Industry Training, Simulation and Education Conference Proceedings, Arlington, Virginia.

Thorndyke, P. R., & Hayes-Roth, B. (1979). The use of schemata in the acquisition and transfer of knowledge. Cognitive Psychology, 11(82-106).

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. Journal of Experimental Psychology: Applied, 3(4), 257-287.

Tomono, T. (2000). Spectacle-type wearable display. Optics Communications, 180, 205-210.

Tranel, D., & Domasio, A. R. (Eds.). (1995). Neurobiological foundations of human memory.

Travis, D. (1991). Effective color displays: Theory and practice.: Academic Press Limited.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. Science, 247, 301-305.

Uemura, T., Arai, Y., & Shimazaki, C. (1980). Eye-head coordination during lateral gaze in normal subjects. Acta Otolaryngol, 90, 192-198.

Vaghan, T. (1998). Multimedia: Making it work. (Fourth ed.): Osborne/McGraw-Hill.

Van Bruggen, J. M., Kirschner, P. A., & Jochems, W. (2002). External representation of argumentation in cscl and the managemnet of cognitive load. learning and Instruction, 12, 121-138.

Van Gerven, P. W. M., Paas, F. G. W. C., Van Merrienboer, J. J. G., & Schmidt, H. G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. learning and Instruction, 12, 87-105.

Ververs, P. M., & Wickens, C. D. (1998). Conformal flight path symbology for head-up displays: Defining the distribution of visual attention in three dimensional space. Aviation Research Lab.

Wade, N. J., & Swanston, M. T. (2001). Visual perception: An introduction. (Second ed.): Psychology Press.

Walker, N., Jones, P., & Mar, H. H. (1983). Encoding processes and the recall of text. Memory and Cognition, 11(3), 257-282.

Waugh, N. C. (1970). Retrieval time in short-term memory. British Journal of Psychology, 61(1), 1-12.

Weeks, K. W., Lyne, P., Mosly, L., & Torrance, C. (2001). The strive for clinical effectiveness in medication dosage calculation problem-solving skills: The role of constructivist learning theory in the design of a computer-based "authentic world" learning environment. Clinical Effectiveness in Nursing, 5, 18-25.

Weintraub, D. J., Haines, R. F., & Randle, R. J. (1993). Head-up display (hud) utility ii: Runway to hud transitions monitoring eye focus and decision times. Proceedings of Human Factors Society 29th Annual Meeting, 615-619.

Weiss, R. E., Knowlton, D. S., & Morrison, G. R. (2002). Principles for using animation in computer-based instruction: Theoretical heuristics for effective design. Computers in Human Behaviour, 18, 465-477.

Welford, A. T. (1968). Fundamentals of skill. London: Methuen.

Wickens, C. D. (1992). Engineering psychology and human performance (2nd ed.): HarperCollins.

Winn, W. (1997). The impact of three-dimensional immersive virtual environments on modern pedagogy., Technical report R-97-15: Human Interface Technology Library.

Winn, W. (2002). Current trends in educational technology research: The study of learning environments. Educational Psychology Review., 14(3), 331-351.

Yanagihara, Y., Kakizaki, T., Arakawa, K., & Isoda, Y. (1998). A multimodal teaching advisor for sensor-enhanced robotic systems in manufacturing. Robotics and Computer-Integrated Manufacturing, 14, 263-273.

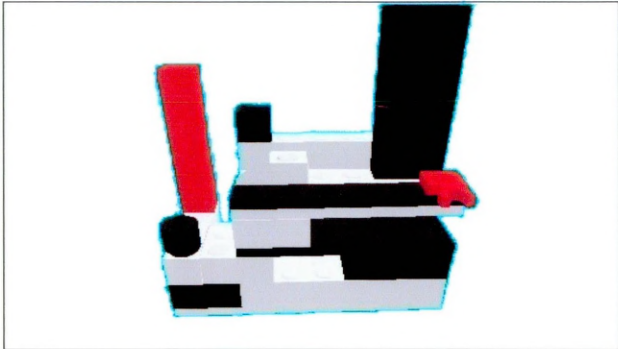
Yerkes, R. M., & Dodson, J. D. (1997). The relation of stimulus to rapidity of habit formation. Nature, 84, 495-482.

Appendices

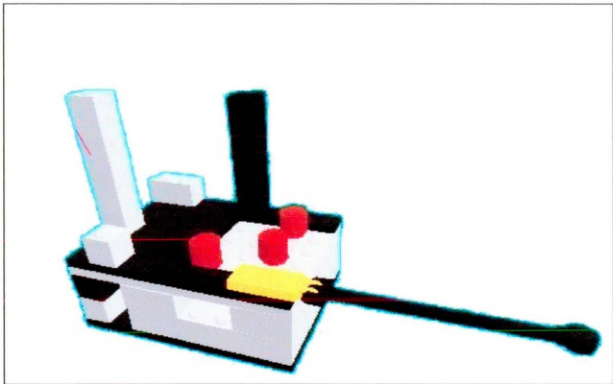
<i>Appendix 3.1: Lego Models used in Pilot Study.....</i>	<i>302</i>
<i>Appendix 3.2: Paper instructions for assembling models in pilot study.</i>	<i>303</i>
<i>Appendix 3.3: Screen shot of animation used as instructions for the Desktop and Head-mounted display conditions in the pilot study.....</i>	<i>304</i>
<i>Appendix 3.4: Post Test Interview for Pilot Study.....</i>	<i>305</i>
<i>Appendix 3.5 Pilot Study: Qualitative Results from semi-structured feedback questionnaire.</i>	<i>307</i>
<i>Appendix 3.6: Counterbalancing for Experiment 1.....</i>	<i>309</i>
<i>Appendix 3.7: Post-test Interview for Experiment 1.1.....</i>	<i>310</i>
<i>Appendix 3.8: Quantitative Results for Experiment 1.1</i>	<i>312</i>
<i>Appendix 3.9: Quantitative Results for Experiment 1.2.</i>	<i>313</i>
<i>Appendix 3.10: Quantitative Results for Experiment 1. 3.</i>	<i>314</i>
 <i>Appendix 7.1: Post-test semi-structured interview for pilot study.....</i>	 <i>315</i>
<i>Appendix 7.2: Feedback from the Semi-Structured Interviews for the pilot study.</i>	<i>317</i>
<i>Appendix 7. 3: Post-test semi-structured interviews for the video alone and the video plus voice conditions in Experiment 2.1.</i>	<i>319</i>
<i>Appendix 7.4: Quantitative results for mean completion times in experiment 2.2.....</i>	<i>321</i>
<i>Appendix 7.5: Quantitative results for mean frequency access in experiment 2.2.....</i>	<i>322</i>

Appendix 3.1:Lego Models used in Pilot Study

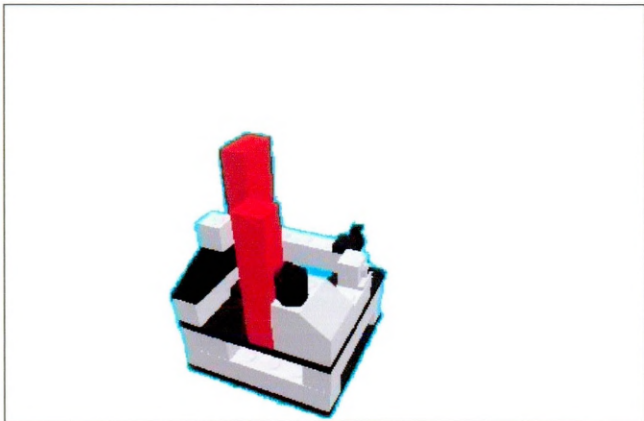
Model A



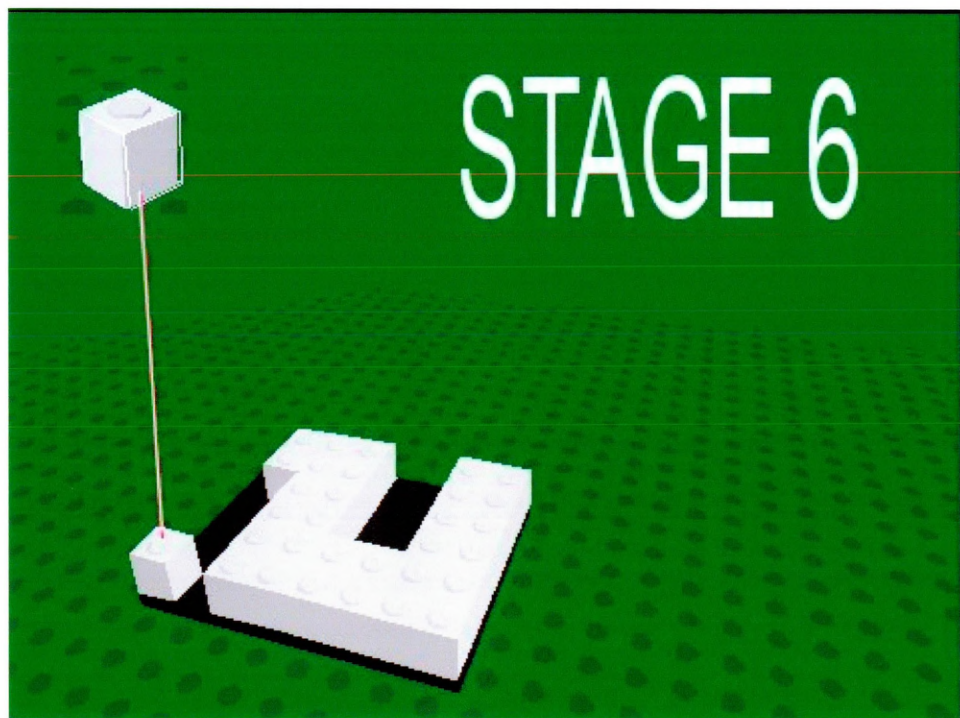
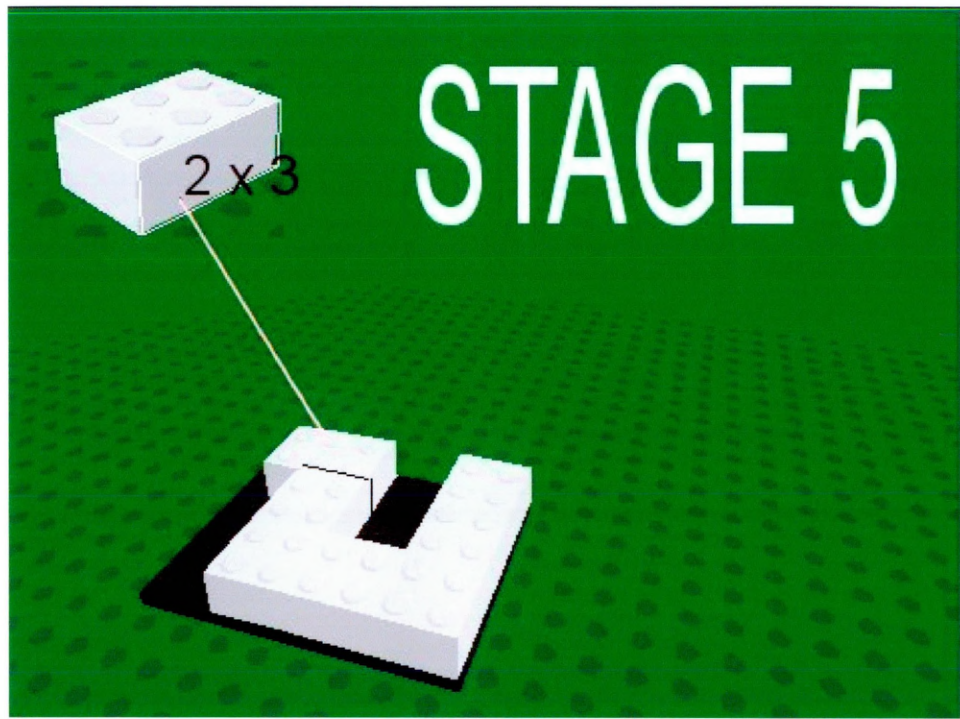
Model B



Model C

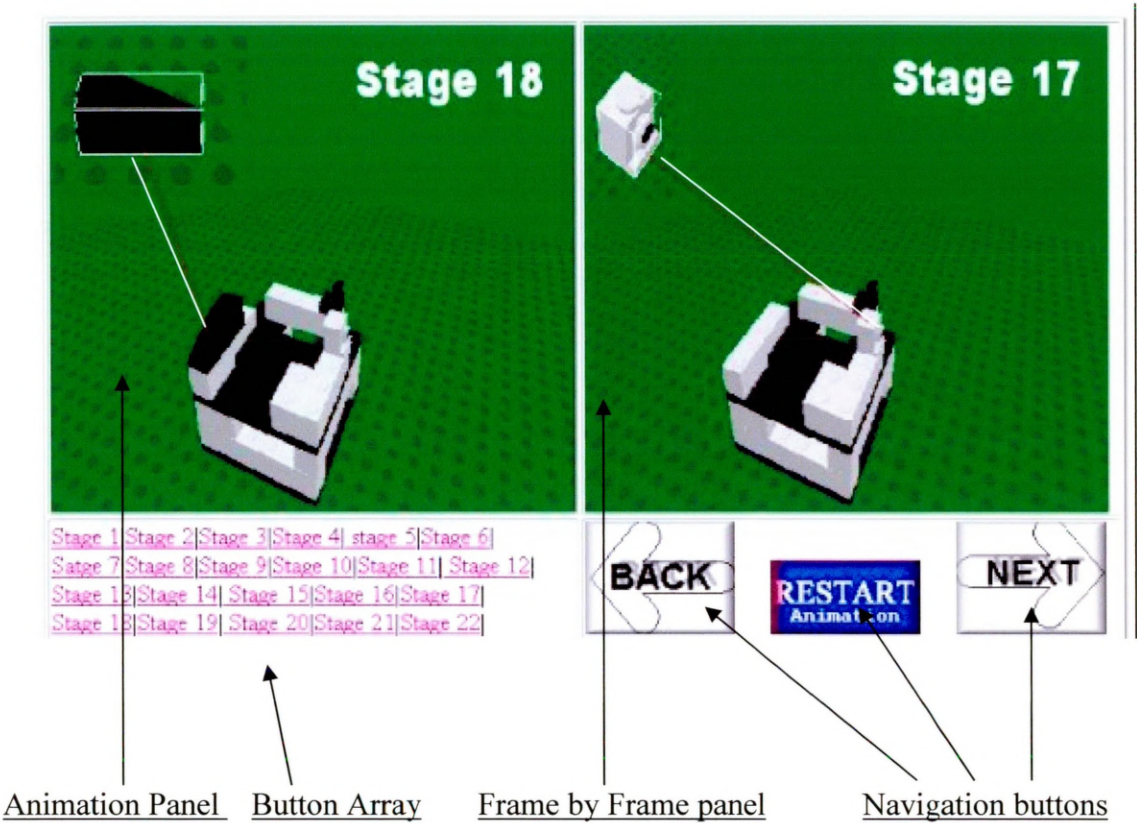


Appendix 3.2: Paper instructions for assembling models in pilot study.



The Paper instructions took the form of a booklet that had two pictures of the assembly steps on each page in colour as above.

Appendix 3.3: Screen shot of animation used as instructions for the Desktop and Head-mounted display conditions in the pilot study.



These instructions comprised of a set of unpublished web pages created using Liquid FX v 3.1 HTML editor. The stages of construction of the models were made using the Lego Creator programme. Then each stage was saved as a jpeg file using the screen capture facility of Paint Shop Pro v.6. The jpeg files were then turned into animations. The animation was made using Animation Shop v.2. that ran at a speed of one frame every 4.5 seconds. The right hand panel contained the instructions in a frame-by-frame format. The left-hand panel contained a full animation of the model beneath the animation are links to each stage in the animation. The current piece to be used is in the left hand corner of each panel, with an arrow indicating its position on the model. The blue “Restart Animation” button starts the animation from the beginning.

Appendix 3.4: Post Test Interview for Pilot Study.

Subject : Age : Gender:

- 1) Which figure did you find easiest/hardest to build?

- 2) Which of the instruction formats, paper, desktop or Head-mounted display did you find easiest to use?

- 3) In the desktop and head-mounted display conditions did you find the option of the two types of instructions:- animation and still frame useful?

- 4) In the desktop and head-mounted display conditions, which instructional format, did you use the most, animation or still frame?

- 5) In the desktop and head-mounted display conditions did you find the speed of the animation either too slow or just right?

6) In the desktop and head-mounted display conditions did you find the control system easy to use?

7) Do you have any further comments about the test?

Appendix 3.5 Pilot Study: Qualitative Results from semi-structured feedback questionnaire.

The first question asked participants which model they found the easiest to build following the instructions. Four participants found model B the easiest to build and one found model C the easiest the other participant thought models A and B were roughly the same but model C harder.

The second question asked which model they found the hardest to build following the instructions. Four participants found model C the hardest to build. However one participant found model B the hardest to construct and another thought model A was the hardest.

The third question asked participants which instruction format they found easiest to use. All the participants except one found the paper instructions easiest to use. One participant reported that they thought the desktop instructions were the easiest.

The fourth question was about the interface on the interaction programme used in the desktop and head-mounted display conditions. The participants were asked which of features of the programme they used the most. All the participants used the animated instructions to construct the models in the all the conditions. No one used the step-by-step still frame feature.

The fifth question asked about the speed of the animation. All the participants except one found the animation too fast at the beginning of the experiment but once they had

adjusted to the speed, they found the animation too slow. One participant thought it was too fast all the time.

The sixth question asked if the participants thought the control system for the animation was easy to use. Five participants reported that they had no difficulty with the controls.

After answering the set questions the participants were invited to give general comments about the experiment. One participant experienced eyestrain using the head-mounted because they were forced to concentrate more and stare into the screen and count the nodules on the play bricks. They had to scan the image to take in all the instructions. Another participant found the head-mounted display uncomfortable at first but got used it eventually. This participant had to put the head-mounted right up to their eyes to see the instruction programme. They felt that the image should be sharper. Two participants complained about the poor image on the head-mounted display compared to the desktop, even though both resolutions were the same, 800x600. They experienced more difficulty in seeing the size and the orientation of the bricks in the head-mounted display than in the desktop, especially the black pieces. One participant found the instructions as seen through the head-mounted display to be too small and had to keep adjusting the display to see the whole screen. One participant suggested that the links to each stage should be put down the side rather than cluttered at the bottom since this made it difficult to pick out a particular link to a stage.

Appendix 3.6: Counterbalancing for Experiment 1. 1

Subject	Presentation Order of Conditions for Each Subject.		
	First	Second	Third
1	Alpha Paper	Beta Desktop	Gamma Micro
2	Gamma Paper	Alpha Desktop	Beta Micro
3	Beta Paper	Gamma Desktop	Alpha Micro
4	Alpha Paper	Beta Micro	Gamma Desktop
5	Beta Paper	Alpha Desktop	Gamma Micro
6	Gamma Paper	Beta Desktop	Alpha Micro
7	Alpha Desktop	Beta Micro	Gamma Paper
8	Gamma Desktop	Alpha Micro	Beta Paper
9	Gamma Micro	Beta Desktop	Alpha Paper
10	Alpha Desktop	Gamma Micro	Beta Paper
11	Beta Desktop	Alpha Micro	Gamma Paper
12	Gamma Desktop	Beta Micro	Alpha Paper
13	Alpha Micro	Beta Paper	Gamma Desktop
14	Gamma Micro	Alpha Paper	Beta Desktop
15	Beta Micro	Alpha Desktop	Gamma Paper
16	Alpha Micro	Gamma Paper	Beta Desktop
17	Beta Micro	Alpha Paper	Gamma Desktop
18	Gamma Micro	Beta Paper	Alpha Desktop
19	Alpha Paper	Beta Desktop	Gamma Micro
20	Gamma Paper	Alpha Desktop	Beta Micro
21	Beta Paper	Alpha Micro	Gamma Desktop
22	Alpha Paper	Gamma Desktop	Beta Micro
23	Beta Paper	Alpha Desktop	Gamma Micro
24	Gamma Paper	Alpha Micro	Beta Desktop
25	Gamma Desktop	Beta Paper	Alpha Micro
26	Alpha Desktop	Beta Paper	Gamma Micro
27	Gamma Desktop	Alpha Paper	Beta Micro
28	Beta Desktop	Gamma Paper	Alpha Micro
29	Alpha Desktop	Gamma Paper	Beta Micro
30	Beta Desktop	Alpha Paper	Gamma Micro

Appendix 3.7: Post-test Interview for Experiment 1.1

Subject : Age:

M or F

1) Which figure did you find hardest to build?

2) Which of the instruction formats, paper, desktop or head-mounted display did you find easiest/hardest to use?

3) In the desktop and head-mounted display conditions did you find the speed of the animation either too slow or just right?

4) In the desktop and head-mounted display conditions did you find the control system easy to use?

5) Did you have any difficulty in identifying the size of the bricks.

6) Do you wear glasses or contact lenses.

7) Do you have any further comments about the test? For Example the head-mounted display condition?

Appendix 3.8: Quantitative Results for Experiment 1.1

A one-way analysis of variance for repeated measures by Instruction (paper, desktop and HMD) for mean completion times.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	2930.751	2	1465.375	.962	N.S.
Error	60920.396	40	1523.01		

A one-way analysis of variance for repeated measures by Instruction (paper, desktop and HMD) for mean error rates.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	20.667	2	10.333	2.431	N.S.
Error	170.000	40	4.250		

Appendix 3.9: Quantitative Results for Experiment 1.2.

A one-way analysis of variance for repeated measures by Instruction (Glasston, Albatech and desktop) for mean completion times.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	34079.874	2	17039.937	5.373	P<0.05
Error	190298.353	60	3171.639		

A one-way analysis of variance for repeated measures by Instruction (Glasstron, Albatech and desktop) for mean error rates.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	63.742	2	31.871	7.093	P<0.05
Error	269.591	60	4.493		

Appendix 3.10: Quantitative Results for Experiment 3.1

A one-way analysis of variance for repeated measures by Instruction (See Through Glasston, non see through Glasstron and desktop) for mean completion times.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	672.667	2	336.333	.641	N.S.
Error	27278.667	52	524.590		

A one way analysis of variance for repeated measures by Instruction (See Through Glasstron , non-see through Glasstron and desktop) for mean error rates.

Source	SS	df	MS	F-Value	Sig Level
Instruction Type	12.247	2	6.123	2.246	N.S.
Error	141.753	52	2.726		

Appendix 7.1: Post-test semi-structured interview for pilot study.

Post Test Interview Pilot4

Subject : Male/Female

Have you Used a HMD before?

1) Which figure did you find easiest/hardest to build?

2) Did you find the speed of the video instructions either too slow too fast or just right?

3) Did you find the control system easy to use?

4)Did you identify the size of the brick by the verbal description or by its video representation or both?

5) Did you find switching attention between the task and the instructions easy or hard?

--

6) Do you have any further comments about the experiment?

--

Appendix 7.2: Feedback from the Semi-Structured Interviews for the pilot study.

The four participants in the pilot study were given a post-test semi structured interview concerning various aspects of the study. There were six closed questions regarding the mechanics of the experiment and an opened ended question asking the participants to give their views on the experiment in general.

Question One. This asked the participants if they had used a head-mounted display before. Three had used a head-mounted display and one had not.

Question Two. The second question asked whether the participants had used Lego before. All participants had previously used Lego.

Question Three. The third question asked participants how familiar they were with car parts. Two reported being very unfamiliar with car parts and the other two intimated that they were very unfamiliar with the parts of cars.

Question Four. This question asked about the speed of the demonstration. When asked if they thought the demonstration was too slow, too quick or just right, all the participants felt that it was just right.

Question Five. This question was different for the two groups. The demonstration only group was asked which multimedia element did they pay more attention to whilst watching the video, visual instructions, verbal instructions or both verbal and visual instructions. Both participants reported that they paid more attention to the visual instructions. The participants in the video only condition were asked if they would

have preferred verbal instructions on the video demonstration. One said they would have preferred verbal instructions and the other said they did not know.

Question Six. This was an open question that asked participants if they wanted to make any further comments about the experiment. One participant thought the assembly was about the right length, any longer would have caused difficulties. Another participant thought the demonstration looked a bit blurry but they could see the screen clearly enough. The third participant thought the speed was good and the task presented no problem. The fourth found it difficult to remember some sections of the assembly compared to other sections.

Appendix 7. 3: Post-test semi-structured interviews for the video alone and the video plus voice conditions in Experiment 4.

Post Test Interview Experiment 2.1 Video Only

Subject : Male/Female

Have you Used a HMD before?

Have you used Lego before?

How familiar are you with the names of car parts.

Very unfamiliar familiar Unfamiliar

Was the pace of the instructional video :

Too fast Too slow Just Right

Would you have preferred some verbal instructions on the video.

Yes No Don't no

Could you list what you liked or disliked about the video

Post Test Interview Experiment 2.1 Video Plus Sound

Subject : Male/Female

Have you Used a HMD before?

Have you used lego before?

How familiar are you with the names of car parts.

Very unfamiliar familiar Unfamiliar

Was the pace of the instructional video :

Too fast Too slow Just Right

Whilst watching the instructional Video did you pay more attention to the:

Visual instructions Verbal instructions

Both verbal and visual

Could you list what you liked or disliked about the video

Appendix 7.4: Quantitative results for completion times in experiment 2.2

A one-way analysis of variance by type of video demonstration (Voice, text, voice plus text and no voice or text) in the initial phase.

Source	SS	df	MS	F-Value	Sig. Level
Between Groups	143845.4	3	47984.457	7.602	p< 0.01
Within groups	346920.9	55	6037.653		
Total	490766.3	58			

A one-way analysis of variance by type of video demonstration (Voice, text, voice plus text and no voice or text) in the criterion phase.

Source	SS	df	MS	F-Value	Sig. Level
Between groups	17634.020	3	5878.007	2.720	N.S.
Within Groups	118848.9	55	2160.889		
Total	136482.9	58			

A one-way analysis of variance by type of video demonstration (Voice, text, voice plus text and no voice or text) in the retention phase.

Source	SS	df	MS	F-Value	Sig. Level
Between groups	94428.7	3	31476.233	7.125	P<0.01
Within groups	242988.1	55	4417.966		
Total	33741.6	58			

Appendix 7.5: Quantitative results for mean frequency access in experiment 2.2

A one-way analysis of variance by type of video demonstration (Voice, text, voice plus text and no voice or text) in the initial phase.

Source	SS	df	MS	F-Value	Sig. Level
Between Groups	144.766	3	48.255	9.124	p< 0.01
Within groups	290.895	55	5.289		
Total	435.661	58			

A one-way analysis of variance by type of video demonstration (Voice, text, voice plus text and no voice or text) in the retention phase.

Source	SS	df	MS	F-Value	Sig. Level
Between groups	74.526	3	5878.007	2.720	P<0.05
Within Groups	270.457	55	4.917		
Total	344.938	58			